```
Db   878 KVIRTPFYTNRQTCQDWLTRIRLSIMRVGLLAGQPAVTVRHGFDLLTEMKTISLSQGNEL 937
         :    |     :   |:  |
Qy   668 --PTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELRIQATW------------YEKLH 714

Db   938 EVTIMVVVEALC----ELHCPEAIGGIAVWSSSIVGK---NLLWINSVAQQAEGRFEKAS 990
         : |:  |   :  |    : : |
Qy   715 -EWEDALVA-------------YDKKMDT------NKDDPELMLGRMRCLEALGEWGQLHQQC 757

Db   991 VEYQEHLCAMTGVDCCISSFDKSVLTLANAGHNSASPKHSLN----------GESRKTV 1039
         : : |      |:       |        :
Qy   758 CEKWTLVNDETQAKMARMAAAAWGLGWDSMEEYTCMIPRDTHD---------------GAF 805

Db  1040 LSKPTDSSPEVINYLGNKACEFYISIADWAAVQEWQNAI----HDLKKSTSSTSLNLKAD 1095
         :  :  : |             |:    |
Qy   806 YRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEEVIQY 865

Db  1096 FNYIKSLSFSEGKFVECTEQLELLPGENINILAGGSKEKI----------DMKKLLPN 1144
         :|  |              |             ::::|
Qy   866 KLVPERREIIRQIWERLQGCQRIV----------------------EDWQKI----LM 898

Db  1145 MLSPDPRELQKSIEVQLLRSSVCLATALNPIEQDQKWQSITENVVKYLKQTSRIAIGPLR 1204
         :|  |        |     |:
Qy   899 VRSLVVS-----------------------------------PHEDMRTW---- 913

Db  1205 LSTIITVSQSLPVLSTLQLYCSSALENTVSNRLSTEDCLIPLFSEALRSCKQHDVRPWMQA 1264
         : | |
Qy   914 LKY---------------------ASLCGKSGRLALAHKTLV---LLLGVD 940

Db  1265 LRYTMYQNQLLEKIKEQTVPIRSHLMELGLITAAKFARKRGNVSLATRLLACCSEVQLGKT 1324
         :|           : |:|       :
Qy   941 PSRQ--------------------------LD-----HPLPTVHP--------Q 955

Db  1325 TTAQDLVQHFKKLSTQGQVDEKWGPELDIEKTKLLYTAGQSTHAMEMLSSCAISPCKSVK 1384
         :|
Qy   956 VTYAYMKNM-----------WK---SARKIDAFQEMQHF-------------VQT 983

Db  1385 AEYAVAKSILITLAKWIQAEWKEISGQLKQVYRAQHCQNFTGLSTLSKNILTLIELPSVNT 1444
         :::|:       :     |
Qy   984 MQQQAQHAIATEDDQQHKQELHKLMARCFLKLGEWQL--NLQGINESTIPKVLQYYSAAT 1040

Db  1445 MEEEYPR-IESESTVH--------------IGVGEPDFILGQLIYHLSSVQAPEVAKSWAALA 1491
         :::|:  :       |
Qy  1041 EHDRSWYKAWHAWAVNF--------EAVLHYKIQNQARDEKKKLRHASGANITN 1087

Db  1492 ----SWAYRWGRKKVVDNASQGEGVRLLPREKSEVQNLLPDTITEBEKERIYGILGQAVCR 1547
         :||||::     :        ::::|
Qy  1088 ATTAATTAATATTTASTEGSNSESE------AESTENSPTPSPLQKKVTEDLSKT---- 1136

Db  1548 PAGIQDEDITLQITESED--NEEDDMVDVIWRQLISSCPWLSELDBSATEGVIKVWRKVV 1605
         :: |
Qy  1137 --LLMYTVPAVQGFFRSISLSRGNNLQD-----------------TLRVLTLW 1170

Db  1606 DRIFSLYKL-SCSAYFTFLKLNAGQIPLDEDDPRLHLSHRVEQSTDDMIVMATLRLLRLL 1664
         :   :  |  : |           :   |:|
Qy  1171 FDYGHWEDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQ 1230

Db  1665 VK--HAGELRQYLEHGLETTPTAPWRGIIPQLFSRLNHPEVYVRQSICNLLCRVAQDSPH 1722
         :  :  :  : | :   :   |
Qy  1231 ALIYPLTVASKSTTTARHNAANK-----------ILKNMCE-----HSN--- 1263

Db  1723 LIIYPAIVGTISLSSESQASGNKFSTAIPTLLGNIQSEELLVSBCEGGSPPASQDSNKDE 1782
         :::||           :
Qy  1264 ------------------TLVQQAMMVSEELIRVAILWHENWHEGL----- 1291

Db  1783 PKSGLNEDQAMMQDCYSKIVDKLSSANPIMVLQVQMIVAELRRVTVLWDELWLGVLLQQH 1842
         :: |           :    |: |           :|    |:      ::|
Qy  1292 ---EEASRL---------------------YFGERN---VKGMPEVLEPLHAMMERGPQT 1324

Db  1843 MVLRRIQQLEDEVKRVQNNNTLRKEEKIAIMRRHTALMKPIVFALEHVRSITAAPAET 1902
         :||                : :    :          |: |||::   |:
Qy  1325 LKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS-LE 1382


Db  1903 PHEKWFQDNYGDAIENALE--KLKTPLNPAKPGSSWIPPKEIMLSLQQRAQKRASYILR 1959
         :||||         :::|:|
Qy  1383 LQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEF 1442

Db  1960 LEEISPWLAAMNTNEIALPGBVSA-RDTVTIHSVGGTITILPTKTKPKKLLFLGSDGKSY 2018
         :::|       ||:||   : || | |         |:|:|     |||:||:|||||||
Qy  1443 VFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGNVPH 1502

Db  2019 PYLFKGLEDLHLDERIMQFLSIVNTMFATINRQETPRFHARHYSVTPLGTRSGLIQWVDG 2078
         :|||||||||:||||||||:||||||||||||||||||||:||||||||||||||:|||
Qy  1503 CDTLHALIRDYREKKKIL-----------------------------LNI 1523

Db  2079 ATPLFGLYKRWQQREAALQAQKAQDSYQTPQNPGIVPRPSELYYSKIGPALKTVGLLSIDV 2138
         :  ||:        ::| |          :    : :   :
Qy  1524 EHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDLLAKLLWLKSPSSEVWFDRRTNYTRS 1583

Db  2139 SRR----DWP-LHVMK--AVLELMEATPPNLLAKELMSSCTTPDEWWRVTQSYARS 2188
         :||     |   |::::|
Qy  1584 LAVMSMVGYIlGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLTRML 1643

Db  2189 TAVMSMVGYIIGLGDRHLDNVLIDMTTGEVVHIDYNVCFEKGKSL-RVPEKVPFRMTQNI 2247
         :||||||||||||||:||:|::|:|:|||:|:||:|::||:||:| ||||||||:||:|
Qy  1644 TNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLRAFVYDPLLNWRL------ 1691

Db  2248 ETALGVTGVBGVFRLSCEQVLHIMRRGRETLLTLLEAFVYDPLVDWTAGGBAGPAGAVTG 2307
         :|    :  |:    ||       :| | ::| |||||||||
Qy  1692 --MDTNTKGNKRSRTR--TDSYSAGQSVEI 1717

Db  2308 GGGQQAESKQSKREMEREITRSLFSSRVAEI 2338
         :|                    |||:|||
```

Search completed: March 2, 2004, 20:10:01
Job time : 70 secs

```
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 3529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-165-216-8

Query Match             7.7%;  Score 726;  DB 15;  Length 3529;
Best Local Similarity  18.6%;  Pred. No. 4.2e-52;
Matches  453;  Conservative 337;  Mismatches 337;  Indels 898;  Gaps 88;

Qy      59 LATIGELAQV---------SGLEMRKWVDELFIIIMDNLQDSSLLAK--RQVALWTLGQLVAST 111
                                                    : :  : :|| : ||||
Db      34 LATVKQLKEFIQQPENKIVLVKQLDNILAAVHDVLNESSKLLQELRQEGACCLGLLCASL 93
                                                                                 :: ::|
Qy     112 GYVVEPYRK----------------------------------------------YPTLLEVLLNFLKT-EQ 136
              ||::| :||
Db      94 SYEAEKIFKWIFSKFSSSAKDEVKLLYLCATYKALETVGEKKAFSSVNQLVMTLQSILE 153
                                    :: ::::|| |
Qy     137 NQGTR--------------------------------------------------------RRAIRVL-----------GLLGA 153
              :: |
Db     154 NVDTPELLCKCVKCLLVARCYPHIFSTNPRDTVDIIVGWIIDHTQKPSLTQQVSGWLQS 213
                                        :: ||
Qy     154 LDPY-------------------------------KHKVNIGMIDQSRDASAVSLBSKSQDSDYSTSEMLVNNGNLPLDEFY 207
              :|: ::
Db     214 LBPFWVADLLAPSTLLGQFLEDMEAYAEDLSHVASGES--------------------------VDBDV 256
                                                        : : :::::::
Qy     208 PAVS----------MVALMRIFRDQSLSHHHTMVVQAITFIFKSLG----------------------LKC 246
              ||| :
Db     257 PPPSVSLPKLAALLRVF----------------------STVVRSIGERFSPIRGPPITBAYVTVIDVLYRVMRC 307
                                                                :::::|
Qy     247 VGFLPGVM--PTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMREFWVM 304
              : : :::
Db     308 VTAANQVFFSEAVLTAANECVGVLLGSLDPSMTIHCDMVITYGLDQLENCQTCGTDYII- 366
                :: :|
Qy     305 NTSIQSTITILLIEQIVVALGGEF--KLYLPQ-----------------------LIPEML--RVFMHD 345
                                        : : : |
Db     367 --SVLNLLTLIVEQINTKLPSSFVEKLFIPSSKLLFLRYHKEKEVVAVAHAVVQAVLSLK 424
                          :||:
Qy     346 NSPGRIVSIKLLAAIQLFGANLDDYLHLLPP----------------------------IVKLF 382
                : :: |||
Db     425 NIPVLETAYKLI--GEMTCALNNLLHSLQLPEACSEIKHEAFKNHVFNVDNAKFVVK-F 481
                                    : :::
Qy     383 DAPEAPLPSRKAALETVDRLITESL------------------------------DFTDVASRIIHPI 419
                :|:
Db     482 DL--------------------SALTTIGNAKNSLIGMWALSPTVFALLSKNLMIVHSDLAVHFPAIQYAV 532
                                                      : ::
Qy     420 VRPTI-----------------DQSPELRSTAMDTLSLSVFPQLGKKYQIFIPMVNKVLVRHR 463
                :: :
Db     533 LYTLYSHCTRHDHFISSSLSSASPSLFDGAV--ISTVTTATKGHFSIILNLLGILLKKDN 590
                                              :::|
Qy     464 INHQRYDVLI-------CRIVKGYTLADE------------------------------- 485
                : ::
Db     591 LNQDTPKLLMTWALEAAVLMEKSETYAFLPSLPSFIKFCKGLLANTLLVEDVNICLQACSS 650
                                                                            :
Qy     486 -----BED-----------PLIYQHRMLRSGCQDALASGPVET---------------------- 512
                ||:
Db     651 LHALSSSLPDLLQRCVDVCRQLVHSGTRIRQAFGKLLKSIFLPDVVLSNNNHTEIQEIS 710
                  :|:
Qy     513 ---GPMKKLHVSTINLQK------------------------AWGAARRVSKDDWLEWLRRISLSELLKDSSSPSL 560
                :: :::|
Db     711 LALRSHWSKAPSNTFHPQDFSDVISFIIYGNSHRTGKDNWLERLFYSCQRLDKRDQSTIP 770
                                        : :|
Qy     561 RSC-------WA------LAQAVNPM--ARDLFNAAFVSCWSELNEDQQDELIRS 600
                ::
Db     771 RNLLKTDAVLWQWAIWEAAQFTVLSKLRTPLGRAQDTTQTI------EGIIRS 817
                                                                :::||||
Qy     601 IELALTS--QDIAEVT----------------------QTLLNLABFMEHSDKG----------P 631
                :|: ::::
Db     818 LAAHTLNPDQDVSQWTTADNDBGHGNNQLRIVLLLQYLENLEKLMYNAYEGCANALTSPP 877
                                                    ::: ::|
Qy     632 LPLRD-----------DNGI----------------------VLLGERAAKCR-AYAKALHYKELEFQKG--- 667
```

```
Qy    1042 HDRSWYKAWHAWAVMNF-------------------EBAVLHYKHQNQAARDEKKKLRHASGANITNA 1088
                                                      : |||: |
Db    1484 ---SWAYRWGRKVVDNASQGBGVRILPREKSEVQNLLPDTITEBEKERIYGILGQAVCRP 1540

Qy    1089 TTAATTAAATATTTASTEGSNSESE-------ABSTENSPTPSPLQKKVTEDLSKT--------- 1136
                                                          :: |||:|
Db    1541 AGIQDEDITIQITESED--NBEDMDVIWRQLISSCPWLSELDESATEGVIKVWRKVVD 1598

Qy    1137 --LLMYTVPAVQGFFRSISLSRGNNLQD------------------TLRVLTLWF 1171
                                                                  :|||:|
Db    1599 RIFSLYKL-SCSAYFTFLKLNAGOIPLDEDDPRLHLSHRVEQSTDMIVMATLRLLLLV 1657

Qy    1172 DYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQA 1231
                :||: :
Db    1658 K--HAGELRQYLEHGLETTPTAPWRGIIPQLFSRINHPEVYVRQSICNLLCRVAQDSPHL 1715

Qy    1232 LIYPITVASKSTTTARHNAANK-----------------ILKNMCE------HSN-- 1263
                ::|:| :|||
Db    1716 ILYPAIVGTISLSSSESQASGNKFPSTAIPTLLGNIQGEELLVSECEGGSPPASQDSNKDEP 1775

Qy    1264 ------------------TLVQQAMMVSEELIRVAILWHEMWHEGL---- 1291
                                        ||
Db    1776 KSGLNEDQAMMQDCYSKIVDKLSSANPTMVLQVQMLVAELRRVTVLWDBLWGLVLLQQHM 1835

Qy    1292 -----EBASRL-----------YFFGERN-----VKGMFEVLEPLHAMMERGPQTI 1325
                :|: ::
Db    1836 YVLRRIQQLEDEVKRVQNNNTLRKEEKIAIMREKHTALMKPIVFALEHVRSITAAPAETP 1895

Qy    1326 KETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFPRRISKQLPQLTS--LBL 1383
                : : ::|: :
Db    1896 HEKWFQDNYGDAIENALE--KLKTPLNPAKPGSSWIPFKEIMLSLQQRASYIILRL 1952

Qy    1384 QYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFV 1443
                :: :|| :
Db    1953 EBISPWLAAMTNTEIALPGEVSA-RDTVTIHSVGGTITILPTKPKLLFLGSDGKSYP 2011

Qy    1444 FLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHC 1503
                :|||  ||| :|||
Db    2012 YLFKGLEDLHLDRRIMQFLSIVNTMFATINRQETPFHARHYSVTPLGTRSGLLIQWVDGA 2071

Qy    1504 DTLHALIRDYREKKKL-------------------------------LNIE 1524
                ||||
Db    2072 TPLPFGLYKRWQQREAALQAQKAQDSYGTPQNPGIVPRPSELYYSKIGPALKTVGLSLDVS 2131

Qy    1525 HRIMLRMAPDYDELTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSL 1584
                : :::: : ::|
Db    2132 RR-------DWP-LHVMK--AVLEELMEATPNLLAKELWSSCTTPDEWWRVTQSYARST 2181

Qy    1585 AVNSMVGYILGLGDRHPSNIMLDRLSGKILHIDFGDCFEVAMTREKFPPEKIPPRLRTRMLT 1644
                |||:||||| :|||: |||
Db    2182 AVNSMVGYIIGLGDRHLDNVLIDMTTGEVVHIDNVCFEKGKSL--RVPEKVPFRMTQNIE 2240

Qy    1645 NAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRL------------- 1691
                |||: :| |||
Db    2241 TALGVTVBGVFRLSCBQVLHIMRRGRETLLTLLEAFVYDPLVDWTAGGEARGFAGAVIGG 2300

Qy    1692 --MDTNTKGNRSRSTR--TDSYSAGQSVEI 1717
                ||||: ||| :
Db    2301 GGQQAESKQSKREMEREITRSLFSSRVAEI 2330
```

```
Qy   1620 DCPFEVAMTREKFPEKIPPRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVWAVLE 1679
               ::||||    :  :  : :||  :|| :::|||:|:::||:|||
Db   2513 ETFPEV------PEIVPFRLTHNMVNGMGMGTEGLFRRACEVIMRLMREQRESLMSVLK 2565

Qy   1680 AFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESI 1739
               ::|:||     :||:|||
Db   2566 PFIHDPIVEW--SKPARGSSKGQVN------ETGEVMNEKAKTHVLDI- 2605

Qy   1740 HSPIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLC 1799
                              :::  ||::||  :  :  :::
Db   2606 ------------EQRLQGVIKTFNRVKGL----PLSIEBGHVHYLIQEATDENLLS 2644

Qy   1800 QCYIGWCPF 1808
               |::|||:|
Db   2645 QMYLGWAPY 2653
```

RESULT 14
US-10-165-216-2
; Sequence 2, Application US/10165216
; Publication No. US20030228675A1
; GENERAL INFORMATION:
; APPLICANT: Otterness, Diane M.
; APPLICANT: Abraham, Robert T.
; TITLE OF INVENTION: ATM Related Kinase ATX, Nucleic Acids
; TITLE OF INVENTION: Encoding Same and Methods of Use
; FILE REFERENCE: P-LJ 5222
; CURRENT APPLICATION NUMBER: US/10/165,216
; CURRENT FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3521
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-216-2

Query Match   7.7%; Score 729; DB 15; Length 3521;
Best Local Similarity 18.6%; Pred. No. 2.3e-52;
Matches 451; Conservative 338; Mismatches 745; Indels 896; Gaps 87;

```
Qy     59 LATIGELAQV-----SGLEMRKWVDELFIIIMDMLQDSSLLAK--RQVALWTLGQLVAST 111
          |||  : :   :   |||     :::|| |: :| :|:   :|    ||| ||| | :
Db     26 LATVKQLKEFIQQPENKLVLVKQLDNILAAVHDVLNESSKLLQELRQEGACCLGLLCASL 85

Qy    112 GYVVEPYRK-------------------------------YPTILEVLLNFLKT-EQ 136
                                                   :::::: ||  |:  |
Db     86 SYEABKIFKWIFSKFSSSAKDEVKLLYLCATYKALBTVGEKKAFSSVMQLVMTSLQSILE 145

Qy    137 NQGTR---------------------------------------REAIRVL---------- 153
              |                                            ::
Db    146 NVDTPELLCKCVKCILIVARCYPHIFSINFRDTVDILVGWHIDHTQKPSLTQQVSGWLQS 205

Qy    154 LDPY------KHKVNIGMIDQSRDASAVSLSESKSSQDSSDYSTSEMLVNMGNLPLDBFY 207
          |:||          :   |||
Db    206 LEPFWVADLAFSTTLLGQFLEDMRAYAEDLSHVASGES---------------VDEDV 248

Qy    208 PAVS-----MVALMRIFRDQSLSHHHTMVVQAITFIFKSLG-------------LKC 246
                   |:|   ::| ::|:           :::::
Db    249 PPPSVSLPKLAALLRVF------------STVRSIGERFSPIRGPPITEAVVTDVLYRVMRC 299

Qy    247 VQFLPQVM--PTFLNVIRVCDGAIREPLFQQLGMLVSFVKSHIRPYMDEIVTLMREFWVM 304
                 :  :  :::|:
Db    300 VTANQVFFSEAVITAANCCVGVLLGGSLGLDPSMTIHCDMVTIYGLDQLENCQICGTDYII- 358

Qy    305 NTISIQSTILLIEQIVVALGGEF--KLYLPQ----------------LIPHML--RVFMHD 345
                      :|| :|    ::     :|  :|                |:|:|    :|:
Db    359 --SVLNLLTLIVEQINTKLPSSSVEKLFIPSSKLLFLRYHKEKEVVAVAHAVYQAVLSLK 416

Qy    346 NSPGRIVSIKLLAAIQLFGANLDDYIHLLLPP------------------------IVKLFD 383
                    :||:||||| |   : ||:|:||  |                          |:||||
Db    417 NIPLVLETAYKLI--LGEMTCALNNLLHSLQLPEACSEHIKHEAFKNHVFNVVDNAKFVVIFD 474

Qy    384 APEAPLPSRKAALETVDRLTBSL-------------------------DFTDYASRIIHPIV 420
          |  |  |||   : :|||::|                             |:|||:||||:|
Db    475 L---------SALTTIGNAKNSLIGMWALSPTVFALLSKNLMIVHSDLAVIFPALQYAVL 525

Qy    421 RTL-------------DQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRI 464
           :|             |  :::        :| :| ::  :  :     |:|:|:|
Db    526 YTLYSHCTRHDHFISSSLSSSPSLFDGAV--ISTVTTATKGHFSIINLLGILLKKDNL 583

Qy    465 NHQRYDVLI----------------------CRIVKGYTLADE- 485
          |   :|:||:                       |::|:|:||:
Db    584 NQDTRKLLMTWALEAAVIMKKSETYAPLFSLPSFHKFCKGLIANTLVEDVNICLQACSSL 643

Qy    486 -----EED-------------PLIYQHRMLRSGQGDALASGPVET 512
                :::               |::: :| | |||:|:||| | |:
Db    644 HALSSSLPDLLQRCVDVCRVQLVHSGTRIRQAFGKLLKSIPLDVVLSNNNHFEIQEISL 703

Qy    513 --GPMKKLHVSTINLQK--------------AWGAARRVSKDWLEMLRRISLELLKDSSPSLR 561
            |:|||:|||:||:|:              ||||||||||||:|||||:|||||||||||:|
Db    704 ALRSHMSKAPSNTFHPQDFSDVISFILYGNSHRTGKDNWLERLFYSCQRLDKRDQSTIPR 763

Qy    562 SC-------WA----LAQAVNPM--ARDLFNAAFVSCWSELNEDQQDELIRSI 601
          ||       ||    ||:||||    |||||||||||||||||||||||:||||
Db    764 NLLKTDAVLWQWAIWEAAQFTVLSKLRTPLGRAQDTFQTI-----EGIIRSL 810

Qy    602 ELALTS--QDIAEVT--------------------QTLLNLAEFMEHSDKG-------PL 632
          ||| :    | : :                      |||||||||:||||||:       |:|
Db    811 AAHTLNPDQDVSQWTTADNDEGHGNNQLRIVLLLQYLENLEKLMYNAYBGCANALTSPPK 870

Qy    633 PLRD------------DNGI-------------VILGERAAKCR-AYAKALHYKELEFQKG---- 667
          |:|             |:|              :|||||:||| :|:|:||||::||||||||
Db    871 VIRTFPFVTNRQTCQDWLTRIRLSIMRVGLLAGQPAVTVRHGFDLLTEMKTTSLSQNELE 930

Qy    668 -PTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATW--------YEKLF- 714
           :|::  :| |  :| | :  |  :::| ::|: |:|:|:|:     |||||:|
Db    931 VTIMMVVEALC------ELHCPEAIQGIAVWSSSIVGK----NLLWINSVAQEGRFEKASV 983

Qy    715 EWEDALVA---------YDKKMDT----------NKDDPELMLGRMRCLEALGEWGQLHQQCC 758
          |  : |: |         ::|:|:           |:| |||||:||:||| |:|:|::| ::::
Db    984 EYQEHLCAMTGVDCCISSFDKSVLTIANAGRNSASPKHSLN----------GBSRKTVL 1032

Qy    759 EKWTLVNDET-QAKMARMAAAAWGLGQWDSMEYTCMIPRDTHD----------GAFY 806
          ::  ::| :   ||| ||||||||  ||::| | :|:| :|::            |||
Db   1033 SKPTDSSPEVINYLGNKACECYISIADWAAVQEWQNAI-----HDLKKSTSSTSLNLKADF 1088

Qy    807 RAVIALHQDLFSLLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEVIQYK 866
          :|::    ::|  ::   |:|||||||:||:|:||| :|| |:|:   :   |:|:|:|:|
Db   1089 NVIKSLSSFESGKFVECTEQLELLPGENINLLAGGSKEKI------DMKKLLPNM 1137

Qy    867 LVPERREIIRQIWWERLQGCQRIV---------------EDWQKI-------LMV 899
          | |:  |::|:|| |||:| |||:|               |:|:|:        |::|
Db   1138 LSPDDPRELQKSIEVQLLRSSVCLATAINPIEQDQKWQSITENVVYLKQTSRIAIGPLRL 1197

Qy    900 RSLVVS------------------PHEDMRTW---------L 914
          ::|::|                  |:|::::|           |
Db   1198 STLTVSQSLPVLSTIQLYCSSALENTVSNRLSTEDCIIPLFSEALRSCKQHDVRPWMQAL 1257

Qy    915 KY--------------------------ASLCGKSGRLALAHKTLV---LLLGVDP 941
          :|                          ||:|||:||:|| ||:|:     |||||:|
Db   1258 RYTMYQNQLLEKIKEQTVPIRSHLMELGLTAAKFARKRGNVSLATRLLIAQCSEVQLGKTT 1317

Qy    942 SRQ--------------------HPLPTVHP---------QV 956
          ||:                    ||||:|:|           ::|
Db   1318 TAQDLVGHFKKLSTQGQVDEKWGPELDIEKTKLLYTAGQSTHAMEMLSSCAISFCKSVKA 1377

Qy    957 TYAYMKNM-----------WK---SARKIDAFQHMQHP--------------VQTM 984
          | | :::|            |:    ||::|:|:|:|::                ::::|
Db   1378 EYAVAKSILTLAKWIQAEWKEISGQLKQVYYRAQHQQMFTGLSTLSKNILTLIELPSVNTM 1437

Qy    985 QQQAQHAIATEDQQHKQELHKIMARCPLKLGEWQL----NLQGINESTIPKVLQYSAATE 1041
           :   |: |:::    ||  :: :|  |::|                |::::|||::::
Db   1438 EEEYPR-IESESTVH----------------IGVGEPDFILGQIYHLSSVQAPEVAKSWAALA- 1483
```

```
Qy   1737 ESIHSFIGDGLVKPEALNKKAIQLINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHE 1796
                                        ||||||:|::|:|
Db   2594 DI---------------------------EQRLQGVIKTRNRVTGL------PLSIEBGHVHYLIQEATDEN 2631

Qy   1797 NLCQCYIGWCPF 1808
          |||:|||:|||
Db   2632 LLCQMYLGWTPY 2643
```

RESULT 13
US-10-227-610-2
; Sequence 2, Application US/10227610
; Publication No. US20030108916A1
; GENERAL INFORMATION:
; APPLICANT: Dunphy, W.
; APPLICANT: Kumagai, A.
; APPLICANT: Guo, Z.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL ASSAYS OF XENOPUS ATR
; FILE REFERENCE: CTCH-P01-003
; CURRENT APPLICATION NUMBER: US/10/227,610
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: 60/314,215
; PRIOR FILING DATE: 2001-08-22
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 2654
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-227-610-2

Query Match   7.8%; Score 734.5; DB 14; Length 2654;
Best Local Similarity 20.7%; Pred. No. 4.7e-53;
Matches 387; Conservative 310; Mismatches 681; Indels 491; Gaps 72;

```
Qy    116 EPYRKYPTILEVLLNFLKTEQ-NQGTRREA---------IRVLGLLGALDPYKHKVNIG 164
          |:|:||||::|| |                        |||  ||  | ::|
Db   1100 EHYQQVFSGLSILATYASNDDPYQGPRNFAKPEIMADYLQPKLGIIAFPNMHILSSSIG 1159

Qy    165 MIDQSRDA-SAVSLSESKSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFR--D 221
          :  |  ||  : :::|:|      ::  | |   |   |    |||  :|  :  |
Db   1160 IEDKKMALNSLVSLMKIMGPKHISSVRVKMMTTLRTLGYRYKEFPGLCCSAWDLFVRCLD 1219

Qy    222 QS----LSHHHTMVVQAITFIFKSLGLKCVQFLP-QVMPTFLNVIRVCDGAIREFLFQQ 275
          |        |  |   |: :|  |  :||                     :  || : | |
Db   1220 QAYLGPLLSH---VIVAL---LPLLHIQPKETVAVPRYLIVENRDAVQDFLHE- 1266

Qy    276 LGMLVSFVKSHIRPYMDEIVTLMREF--WVMNTSIQSTIILLI-------EQIVVAL--- 323
          |||||:|:|||  |  |:  | :|        :|   :|  |         || |
Db   1267 ---IYFLPDH--PELKEIQKVLQEYRKEFTKSTDLQTAMQLSIRAIQHENVDVRMHALT 1320

Qy    324 ------GGEFKL------------YLPQLIPHMLRVFMHDNSPGRIV--SIKLLAAIQ 361
                  |  |||           :| ||  |  | :||    |    ::||| |: |
Db   1321 SLKETLYKNQAKLLQYSTDSETVEPVISQIVTVLLIGCQDANPQARLFCGBCLGQLGAID 1380

Qy    362 -----------LFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLIT 403
                     |  ||  || |  | : | |:    |||  | ||
Db   1381 PGRLDFSPSETQGKGFTFVSGVEDSDFAYBLLIBEQTRAFLAYADNVRAQDSAAVAIQELI 1440

Qy    404 ESLD-----------FTDYASRIIHPIVRTLDQSPELRSTAMDTLSSLLVF--QLG 445
          |:|             | ::  | ::|  |:   |::  :  ::  | :::
Db   1441 SIFBCKEGRTDCPGRRLWRRFPEHVQEILBPELNTRYKSSR-KAVNWSRVKKPIYLSKIG 1499

Qy    446 KKYQIFTP----MVNKVLVRHRINHQRYDVLICRI-------VKGYTLADEEBDPLIYQ 493
          ||  |      |||:||    | |:||   |:|         ||  |||  | |||:
Db   1500 NNFADWSAIWAGYLITK--VRHELARRVFS---CCSIMMKHDFKVTIYLL-----PHILV 1549

Qy    494 HRMLRSGQGDALASGPVETGPMKKLFVSTINLQKAWGAARRVSK-DDWLEWLRRLSLELL 552
          ::  | :  |:  :                 ||    | | || ||  :| | | ::
Db   1550 YVLLGCNKEDQ----------------QEVVAEIMAVLKHEDPL--MRRLIQ-DSA 1585

Qy    553 KDSSSPSLRSCWALAQAYNPMARDLPNAAFVSCWSELNEDQQDEL--IRSIELALTSQDI 610
          |::::|: |                                    ::||::  :::|:
Db   2453 TFPDPTSWYNSRSAYCRSTAVMSMVGYILGLGDRHGENILFDSLTGECVFDVFNCLFNKG 2512
```

```
Db   1586 SDLSQLSTQTVFSMLDHLTQWAREKFQA---------------LNAEKTNPKPGTGEPKAVSNEDY 1637

Qy    611 ABVTQTLNLAEFMEHSDKGFPLPLRDDNGIVLLGERAAKCRAYAKAL-HY------KELE 663
          | :|:|:|  | | :   ||  | |  |:| ||  | | | : :: |:         ||:|
Db   1638 GEYQ-----NVTRFLDLIPQDTLAV------------ASFRSKAYTRALMHPESFIMEKKQE 1682

Qy    664 FQKGPTPAILESLLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHBWEDALVAY 723
          :||  |  |    |   |::      | |   |  :  |:|| : : :  |::|:|  ::|
Db   1683 IQEH---LGFLQKLYAAMHEPDGVAGVSAIRKKE-ASLKEQIIEHESIGLLRDATACY 1736

Qy    724 DKKMDTNKDDPELMLGRMRCLEALGEWGLHQQCCKWTLVNDETQ--AKMARMAAAAAW 781
          | |   | |  |    | |:| | | | | ::    |   | |    | | | | |
Db   1737 DRAIQLK----PEEIIHYHGVVKSMLGLGQLQSTVITQVNGILNSRSEWTABLNTYRVEAAM 1793

Qy    782 GLGQWDSMEBYTCMIPRDT----HDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELIA 837
                                  | || :| :|: :|::|  :  |:|  | :: |:
Db   1794 KLSQWDLIVEEYLSADRKSTTWSIRLGQLLLSAKKGERDMFYETLKVV-RAEQIVPLSAAS 1852

Qy    838 MAGESYSRAYGAMVSCHMLSBLEEVIQYKL---------VPERREIIRQIWWERLQ 884
          :| :: |  |          : |  ||                | :|  |: |:|
Db   1853 FERGSYQRGYEYIVRLHMLCELEHSVKMFLQKPSVEPAVDSLNIPARLEMTQNSYRAR-- 1910

Qy    885 GCQRIVEDWQKILMVRSLVVS-----PHEDM--RTWLKYASLCGKSGRLALAHKTLVLLL 937
                                                |: | | : || : :  |:||||
Db   1911 -------EPILAVRRALQTINKRPNHADMIGECWLQSARVARKAGHHQTAYNA---LL 1958

Qy    938 GVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQ 997
                                  | : |   :| ::||   :|  :: | | : ::
Db   1959 NAGESRL--------SELNVERAKWLWS---KGDVHQAL------IVLQKGAEHLFLSSTSA 2002

Qy    998 QHKQELHKLMARCFLKLGEWQLNLQGINEBSTIPKVLQYSAATEHDRSWYKAWHAWAVMN 1057
          ::|                             ||:     |::  | | : | |: :|
Db   2003 PPEQQLI--IHGRAMLLVGRLMEETANFESNA---VMKKYKDVTALLPEW-EDGHFYLAKY 2056

Qy   1058 FEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATTASTEGSNSESEABSTE 1117
          :                                               :  :|:|
Db   2057 YDKLMPWTDNKMEKQGDLIRY--------------------------------2078

Qy   1118 NSPTPSPLQKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTIWFDYG--- 1174
                                   ||     |    |  |   |  ::|:|: ::|
Db   2079 ------------IVLPPGRSLQFGNQYIYQSMPRMLSWLDFGAKV 2112

Qy   1175 -HWP------------DVNEALVBGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIH 1218
                                    | | |  | :  |       |    | :  |
Db   2113 YEWRKAGRADRLQMGNELMKINKVISDHKNQLAPYQFLTAFSQLISRICHSHDEVFAVLM 2172

Qy   1219 QLLTDIGRYHPQALIYPLTVASKSTTARHNAANKLLKNMCEHSNTL--VQQAMMVSEE 1275
          |      | |  |         |:|      |:  |         |  : ||  :
Db   2173 EIVAKVEVAYPQQAMWMTAVSKSSYPMRVNRCKEILEKALHMKPSLGKFIGDATRLTDK 2232

Qy   1276 LIRVAIIWHEMWHEGLEEASRLYFGERNVKGM---FEVLPLHAMMERGPQTLKETSFNQ 1332
                                               |:  | | |  |   |  ::|
Db   2233 L------------LELCNKPVDGNTSTLSMNIHFKMLKKL---------VEETTFSE 2268

Qy   1333 AYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLIYHVPRRISKQLPQLTSLELQYVSPKLLM 1392
                                                              | | :|:
Db   2269 I----LIPLQ-----------------------------------SVMIPTL-- 2286

Qy   1393 CRDLELAVPGTYDPNQPI---IRIQSIAPSLQVITSKQRPRKLILMGSNGHEPVFLLKG 1448
                              |:| |:|     :|:  | | ::   | : |  ::|
Db   2287 KRD--------HADHDPFPGHWAYLSGPDDAVEILPSLQKPKKISLKGSDGKSYIMCKP 2338

Qy   1449 HEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPH-CDTLH 1507
          :| |:|: :|::     |     |   : | | :| |||:| |:   | ||
Db   2339 KDDLRKDCRLMEFNSLINKLRKGABSRRRELHIRTYAVIPLNDBCGIIEWVNNTAGFRN 2398

Qy   1508 ALIRDYREKKLILNIEHR--IMLRMAPDYDHLIMQKEVFBHAVNNTAGDDLAKLLWL 1565
          |     :| : |::|:|      ||                |  | |    | :| :
Db   2399 ILIKLYKEKGIYMGGKELRQCMLPKSAP-----------LQEKLKVFKEALLPHHPPLPHEWFLR 2452
```

```
; SEQ ID NO 39
; LENGTH: 2644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-300-453A-39

Query Match          8.08;  Score 753.5;  DB 14;  Length 2644;
Best Local Similarity  20.3%;  Pred. No. 1e-54;
Matches  428;  Conservative  324;  Mismatches  755;  Indels  605;  Gaps  81;

Qy    32 MEPIIKALILKLKDPDPDNPGVINNVLATIGELAQVSGLEMRKWVDELPIIIM-----  85
         ::  :  : ::|:||                       ::|| :
Db   802 VKAVLGTTLNLMEDPDKDVRVAFSGNIKHIESLDSEDG-----FIKELFVLRMKRAYTH 856

Qy    86 ------DMLQDSSLLAKRQVALWTLGQLVASTGVVVEPYRKYPTLLEVLLNFL--KTEQN 137
               ::| |              |||                ::| :
Db   857 AQISRNNELKDTLILITGDIGRAAKGDLIV----------PFALLHLLHCLLSKSASV  903

Qy   138 QGTRREAIRVL------GLLGALDPYKHKVNIGMIDQSRDASAVSL------------SES 180
          |  :  | |       ::|                       :        :    :::
Db   904 SGAAYTEIRAILVAAKSVKLQSFFSQYKKPICQPLVESLLHSSQMTALPNTPCQNADVRKQD 963

Qy   181 KSSQDSSDYSTSEMLVNMGNLP-LDEPY----------PAVSMVALMRIFRDQS     223
         |  :   |  ::                           |||| |:|| ||||
Db   964 VAHQREMALNTLSEIANVFDFPDLNRPLTRTLQVLLPDLAAKASPAAS--ALIRTLGKQL 1021

Qy   224 LSHHHTMVQAITFIFKSLGLKCVQ--------FLPQVMPTFLNVIRV               263
          : |:          |:||||                :|    :|::
Db  1022 NVNRREILINNPKYIFSHIVCSCSKDELERALHVLKNETEIELGSLLRQDFQGLHNELLL 1081

Qy   264 CDGAIREFLFQQLGMLVSFVKSHIRPYMD--EIVT--LMRE------FWVMN---     305
          ||  | | | ||                    :|                |:|
Db  1082 RIGEHYQQVFNGLSILASFASSD-DFYCGPRDIISPELMADYLQPKLLGILAFFNMQLLS 1140

Qy   306 ----------TSIQSTIILL----IEQIVAL----GGEFKLYLPQLIPHMLRVFM       343
                      ::|  :|         |::|       ||         :   | | :
Db  1141 SSVGIEDKKMALNSLMSLMKLMGPGHVSSVRVQMTTLRTGLRFKDDFPPELCCRAWDCFV 1200

Qy   344 H--DNS-PGRIVSIKLLAAAIQLFGAN------------LDDYLHLLLPPIVKLF       382
                  | |:|  :|| |  :||                :||||  |::|||
Db  1201 RCLDHACLGSLLSHVIVALLPLIHIQPKETAAIFHYLIIENRDAVQDFLH------BIY 1253

Qy   383 DAPEAP-LPSRKAALETVDRLTESLDFTDYASRIIHPIVRTLDQSPELRSTAMDTLSSLV 441
              | |:|||          :|               :|||                :|||
Db  1254 FLPDHPELKKIKAVLQEYRK--ETSESTDLQTTLQLSMKAIQHENVDVRIHALTSLKETL 1311

Qy   442 FQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIV---------------         477
          :: | |                       ::                 :
Db  1312 YKNQEK----------------LIKYATDSEVVEPIISQLVTVLLKGCQDANSQARLLCGECLGE 1360

Qy   478 ----------------KGYTLADEEEDP------------                     489
                          | :||
Db  1361 LGAIDPGRLDFSTTETGCKDFTFVTGVEDSSFAYGLLMELTRAYLAYADNSRAQDSAAYA 1420

Qy   490 -------LIYCHR-MLRSGQGDAL-ASGPVETGMKKLHVST--INLQKA--WGAARR---V 535
                | ||:  ||::  |     ||||| :|||||                    :
Db  1421 IQELLSIYDCREMETNGPGHQLWRRFPEHVREILEPHLNTRYKSSQKSTDWSGVKKPIYL 1480

Qy   536 SK-----DDW-LEWLRRLSLELLKDSSSPPSLRSCWALAQAYNPMARDLFNAAFVSCWSEL 589
         |:      |  :| |                             ||       |:   |::
Db  1481 SKLGSNFAEWSASWAGYLITKVRHDLASKIFTCCSIMMKGHDFKVTIYILLPHILVYVLLGC 1540

Qy   590 NEDQQDELIRSI--------ELALTSQDIAE----VTQTLL-----NLAEFMEH----   626
         | :| |  : :         :||   ||           ||      | ||::
Db  1541 NQEDQOEVVAEIMAVLKHDDQHTINTQDIASDLCQLSTOTVFSMLDHLTQWARHKFQALK 1600

Qy   627 SDKGPLPLRDDNGI---------------------VLLGERAAKCRAYAKA-LHYKE   661
         : |  :|:                             ||       | |    |:::
Db  1601 AEKCPHSKSNRNKVDSMVSTVDYEDYQSVTRFLDLIPQDTLAVASFRSKAYITRAVMHFES 1660

Qy   662 LEFQKGPTPAILESLLISINNKL----QQPEAAAGVLEYAMKHFGBLEIQATWYEKLHEWE 717
          ::|:|    :| :|  :|       |:|  |||| ::  | | | :|   |::|| ::
Db  1661 FITEKKQN--IQEHLGFLQKLYAAMHEPDGVAGV-SAIRKALEPSIKEQILEHESLGLLR 1716

Qy   718 DALVAYDKKMDTNKDDPELMLGRMRCLEAJGEWGQLHQQC-----CEKWTLVNDETQAK   771
         || :| ||  :  :|||:        | || |  ||            :| ||:|:: :
Db  1717 DATACYDRAIQLEPDQIIHYGVVKSMLGLGQLSVITQVNGVHANRSEWT--DE----     1769

Qy   772 MARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARD-   829
         :    | |  | | | ||| | |||                 |:| ||      |:| |
Db  1770 LNTYVREAAWKLSQWDLVENYLAA-------DGKSTTWSVRLGQLLLSAKKRDITAFVDSL 1823

Qy   830 -LLDAE------LTAMAGE-SYSRAYGAMVSCHMLSELEVIQ--YKLVP--ERREIIRQ   877
          :| |:     |:|: | | |: | | |                 |:||     ||::|
Db  1824 KLVRAEQIVPLSAASPERGSYQRGYBYIVRLHMLCELEHSIKPLFQHSPGDSQEDSLN-  1882

Qy   878 IWWERLQGCQRIVEDWQKILMVRSLNVS-----PHEDM--RTWLKYASLCGKSGRLALAH 930
           | ||  ::  :|  :  |::|| |          |:  |     |: |  :  :  |
Db  1883 -WVARLEMTQNSYRAXEPILALRRALLSLNKRPDYNEMVGECWLQSARVARKAGHHQTAY 1941

Qy   931 KTLVLLLGVDPSRQLDHFLPLPTVHPQVTYAYMKQNWWKSARKIDAFQHMQHFVQTMQQQAQH 990
          | | | |  ||      |         ::   :| :::| ::  ||
Db  1942 NA--LLNAGBSRI--------AELYVERAKWLWSKGDVHQALIVLQKGVSLCFPENE-   1988

Qy   991 AIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQYYSAATEHDRSWYKAW 1050
           :|      ::|:|    ||  |||:      ||:   :  ::   :|   ||: : |
Db  1989 ---TPPBGKNMLIH--GRAMLLVGRFMEETANFESNAIMKKYKDVFACLPE--------  2034

Qy  1051 HAWAVMNFEAVLHY-KIQNQARDEKKKLRHASGANITNATTAATTATTTASTEGSNS    1109
          |       ||| | ::  :                 ||
Db  2035 --WEDGHFYLAKYTDKLMPMVTDNK-------------------                 2057

Qy  1110 ESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVCGFPRSISLSRGNNLQDTLRVLTL 1169
              |      |||:|| : |:     ::|:|    |::|     :   :|||  |:|:|:
Db  2058 ----MEKQ--GDLIRYIVLH----------DVNEALVEGVKAIQIDTWLQVIPQLIARIDTPR 2093

Qy  1170 WPDYG----HWP--------------                                   1210
          ||||     ||
Db  2094 WLDYGTKAYEWBKAGRSDRVQMRNDLGKINKVITEHTNYIIAPYQFLTAFSQLISRICHSH 2153

Qy  1211 PLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTI---VQ  1267
         |: ||       :||  |||  :||  |: :   |        | |:  :||        |
Db  2154 DEVFVLMEIIAKVFLAYPQQAMVMMTAVSKSSYPMVRNRCKEILNKCAIHMKKSLEKPVG 2213

Qy  1268 QAMMVSEELIRVAILWHEMWHEGLEEASRLIYPGERNVKGMFEVLEPIHAMMERGPQTLKE 1327
         :|::|  |:| |: |:|:| || | || |:|||  : | : |  | : :|    | ||:
Db  2214 DATRLTDKLL-----ELCNKPVDGSS-----STLSMSTHFKMLKKL------------VEE 2252

Qy  1328 TSFNQAYGRDLMEAQEWCRKYMKSGNVKDETQAWDLYYHVFRRRISKQLPQLTSLELQYVS 1387
         ::| || |: |:           || :|             |  :: ||:
Db  2253 ATFSEI----LIPLQ---------------------SVMIPLPSI-                2273

Qy  1388 PKLLMCRDLELAVPGTYDNQPI------IRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFV 1443
         |:||  || | ||  | | ||        : ::    | :|:|  |:| | | |||:| :
Db  2274 ------------------------LGTHANHASHEPPGHWAYIAGFDDMVEILALGLQXPKKISLKGSDGKFYI 2323

Qy  1444 FLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHC  1503
         | |||:|:|: ||:: |:|: :| : | |  : | |||| :: |  |:|| | ||  |
Db  2324 MMCKPKDDLRKDCRLMEFNSLINKCLRKDAESRRRELHIRTYAVIPLNDECGIIEWVNNT 2383

Qy  1504 DTLHALI-RDYREKKKILLNIEHRIMLRMAPDYDHLITMQKVEVEHAVNNTAGDDLAKL  1562
         ||: | : | :: :| ::|:  | : ||  |:| ||  :|:|  ::|   :|  ||:|
Db  2384 AGLRPILTKLLYKEKGVYMTGKE-----LRQCMLPKSAALSEKLKVFEFLLPRHPPIPHEW 2439

Qy  1563 LWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDHPSNLMLDRLSGKILHIDF----  1618
         |::|  |:|:|  | | | |:|||||:|||||| ::| | ||:|  ::|:::|
Db  2440 FLRTFPDPTSWYSSRSAYCRSTAVMSMVGYILGLGDRHGENLLFDSLTGBCVHVDFNCLF 2499

Qy  1619 -GDCFEVAMTREKPPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMA  1676
              |: |     ::|: || :|       | ||| || :   |: | :::| |:|::
Db  2500 NKGBTFEV----------PEIVPFFLTHMVNGMGPMGTEGLFRRACEVTMRLMRDQREPLMS 2552

Qy  1677 VLEAFVDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILLDGVELGEPAFKKTGTTVP 1736
         ||: |  |::|   ::|:             |:  :|:::   :       || || :
Db  2553 VLKTFLHDPLVEW----SKPVKGHSKAPLN-----                          2593
```

```
Qy   152 --------GALDP-YGHKVNIGMIDQSRDASAVSLSESKSSQDSDYSTSEMLVNMGN          200
                  :  ::|    ::   :                       :  :|::|
Db   956 TTARSQVCCSGSVTPAWKHLYSPHSSHRYRHVYCLALGCDELRSDAPS-CWSVMIMNMDE    1014

Qy   201 LPLDEFYPAVSMALMRIPRDQSLSHHHTM----------VVQAITPIFKSLGLKCVQF       249
          :|        |::|                        :
Db  1015 MDVPNLLDTTLWV---------ITHHWTCFDDSTKQQAQGVLDALS--KINGHDLVKA      1061

Qy   250 LPQVMPTFLNVIRVCDGAIREPLFQQLG----MLVSFVK----SH--IRPYMDEIVTL       297
          :|        |    :|                |:|::
Db  1062 TAS-MUTFNHLPDLAEHRKRLGLLSQSGLTRQDLFRTFCRRLRHDHPQVIEQALTELLEL    1120

Qy   298 MREFWVMNTSIQSTIILLIEQIVALGGEFKLYLPQLIPEMLRVFMHDN-SPGRI-VSI       354
                  :|::  :|::  |:
Db  1121 LRQH--QDYIQAS--------ALAEQPKSFIPGLTRSVLDCAAKYNGWQPGIMRLCA      1167

Qy   355 KLLAAIQLPGAN------------LD------DYLHLLLPPIVKLFDAPEAPLPSR         392
                                  ::        ::
Db  1168 ECMGIIGCLDSNRMEBSNREQKRFVVSQNFMDFRETADFVCFMLENVLV---------     1215

Qy   393 KAALETVDRLTESLDFTDYAS-------RIIHPIVRILD                          424
          ::|                           ::
Db  1216 KAFLSTTD--TKFLSFLSYAMQELLAKTGIKLAYQSQGQQHETLYKRWRSFNDITKEI      1272

Qy   425 QSPELRSTAMDTLSSLVFQLGKKYQIFIPMVN----KVLVRHRI-NHQRYDVLI-----      473
                 | : :|::  :|     |::|:|
Db  1273 LIPFLTSEFI--VSSGMIGLDTNYPIFHPQKSYETWLKITPVLDLLRNPQNPNSAILFPPL  1330

Qy   474 CRIVKGY-TLADEEEDPLIYQHRMLRSCGQDALASGPVETGPMKKLHVSTINLQKAWGAA     532
          || |     |  |||    :|   |          |:|||   :: |:|||
Db  1331 CRLIKVHDPCVSEFLLPYVVLHAVIGS-----EGPSQTPDESQMLVAEBGSTRAFAHE      1383

Qy   533 RRVSKDDWLEWLRRLSLELI----------KDSSSPSLRSCWALAQAY--NPMAR          575
          ::                          :  ::
Db  1384 KKTSGS---EWRKKVMSELKTHLEYQPPETATHSEKEEVKLCYEVCTFLMTVYYINXLKX    1440

Qy   576 DLFNAA-FVSCWSELNEDQQDELIRSIELALTSQD-----IAEVTQTLLNLAEFMEHSD      628
          :  |         ::                  :
Db  1441 AVFRILDYFKNWLHIKKSQS----------AITKEKDARWCPLIVEEVLSSL---------  1480

Qy   629 KGPLPLRDDNGIVLLGBRAAKCRAYAKALHYKELEFQKGPTPAI----------LESLISINN  681
          :|        :||||                  :||
Db  1481 -----DPE---LVARRAIDCGQYARALLFLEPHIESRRDEAIGDEATRLMRSVHDIYT      1530

Qy   682 KLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDALIVAYDKKMDTNKDDPELMLGRM    741
          ::        ::|           ||  :||::|||    |:||
Db  1531 QIDDPDGLDGI-SACMKDLGFKB-QALSHRKAGRWTAACTWYEIQLAESPDDVNLQLDLL    1588

Qy   742 RCLEALGEWGQLHQQCCEKNTLVN----------DETCAKWARM--AAAAAWGL           783
          ::                      :  :
Db  1589 TCLERSGQYGEXHLLSLGKWYIRSXHCVQTIFXVLLEGIDKIPSSLSKVMPFVLEASWAT    1648

Qy   784 GQWDSMERYTCMIPR---DTHDGAFYRAVLAHQDLFSLACQCIDKARDLLDAELTAMA       839
                       |:    :      :   : : :
Db  1649 GRWQIMEKYLRSYTEGDVTDVFNIGIADALLCLKBGDGERFQELLQAMRDKVASSMALSA    1708

Qy   840 GESYSRAYGAMVSCIMLSELEEVIQYKLVP-ERREIIRQIWERLQGGCQRIVEDWQKILM     898
          |: |::::|:    ::                       :
Db  1709 TSSFRTCHDVMLKCHVLEDLEMIANAEPVEGEGHAPLVKALERRLEVLGAYVSDKQYVLG    1768

Qy   899 VR---------SLVVSPHEDM----RTWLKYASLCGKSGRLALAHKT                  932
          ::                       :|:::|:::|:|| | :
Db  1769 VRRAAMELMRYVSCLISELIELISDKSLGRNSETKRYSSSWLATARLARKSGS--THQS     1825

Qy   933 LVLLIGVDPSRGLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQEMCHFV-QTMQQQAQHA     991
          ::                       :  :
Db  1826 FNAVL--RAQQLGDSSAVIE-----YAKLFYKDGQERKAIQLLQRAIDDLFNDGMMA      1876

Qy   992 IATEDQQHKQELHK-LMARCFLKLGEWQLNLQGINESTIPKVLQYSAATEHDRSWYKA      1049
          |:         |:   |:||  |:|||:
Db  1877 IDTPTSSKSQSHRNLLKARAHLLLAKW---LDSTGQTHAGALRSKFQEAAKTHPQWEKG    1933

Qy  1050 WHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAANTTAATATTTASTEGSNS  1109
                    :|::                :|
Db  1934 ------------HYYLGRHYKKVLESEKALSPDDQSDAYLTGETAKL------          1968

Qy  1110 ESEAESTENSPTPSPLQKKVTEDLSKTILMYTVPAVQGFFRSISLSRGNNLQDTLRVLTL    1169
                                          | :::|:  :||::|| :||
Db  1969 ------------VIENYLRSLNFGIKYVHQTLPRILDL                          1994

Qy  1170 WFDYGEWPDVNEALVEGVKAIQIDTWLQ------------VIPQ                    1201
          | :                                      ::|
Db  1995 WLELGTQVD----APSLGMVTLSAELQSRRTILHELYKGHFNRHLPKMPAYIFYTALPQ    2049

Qy  1202 LIARIDTPRPLVGRLIHQLLTDIGRYHPQALLYPLTVASKSTTARHNAANKLLKQMCEH     1261
          |:|| |: ::  |:|::::||    |  ||
Db  2050 IVARIAHPNQDVFRVLEQMIKVRAHPRQAIWSLF----SFMTGRTNGARR------        2097

Qy  1262 SNTLVCQAMMVSEBLIRVALIWHEMWHHGLEEASRLYFGBRNVKGMFEVLEPLHAMMERG    1321
                                                        :|::
Db  2098 ------------HRGQ------------KVLDDLRAIAKRV                       2114

Qy  1322 PQT----LKE-TSFNQAYGRDLMEAQEWCRK------YMKSGNVKDLTQAWDLYHVFRR    1370
          |:       :   :   ||                    |::|:||| || ||:||:|||
Db  2115 DETGYDLKQLLKMGEKLAEQLLLA---CNKGDFQSNRTVKASITRDLN------FNH     2162

Qy  1371 ISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQ------PIIRIQSIAPSLQVI     1423
          :                       ::                  |:|: :::  |:::
Db  2163 KCTPCPLVVPIE----------TC-LTATLPLTDNTRKHKAPSGDVITIDRFLDDVVL     2211

Qy  1424 TSKQRPRKLTLMGSNGHEPFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQ   1483
          :|: ||||||  |||:| :: :|| |:|||:||:|:|::|||:|  | |||:||||::||
Db  2212 GSLAKPRKLTARGSNGQLYGLLIKPKDDLRTDQRLMEFNGLINRSLKRDTESSKRQLVIR    2271

Qy  1484 RYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQK    1543
           ||| ||  :||:||| ::  || :|::|::::|:|::|  :    |
Db  2272 TYAVTPLNEECGIIEWVDGLKTL------RDILLGIYK--TRNITPNYG------QI     2314

Qy  1544 VEVFEHAVNNTAGDDLAKLLWLKS------PSSEVWFDRRTNYTRSLAV              1586
          :|::::| ::||| ||||||:|                    ::|||:|:|
Db  2315 AELMKQAC--TSDDNLH-LWSRSVLGMFPDVLPEWFISQFPDPSAWFAARLRYTRSCAV    2370

Qy  1587 MSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPFKIPFRLTRMLTNA    1646
          |||||:||||||||| .:: |::  ||::|:||| ||:|| :|| |:|||||| :::|
Db  2371 MSMVGTIILGLGDRHGENVLLEEGNGGVFHVDFNCLFDKGLTFAQ-PEKVPFRLTHNMIAA   2429

Qy  1647 MEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRT    1706
          | : ::||:: .| | |:|::|  :  :|::|||  ||              :|:||||
Db  2430 MGIYRYEGPFRNCSELITLKVLRQQEETIMTILEAPLHDPTLDL------QRTKKRT      2479

Qy  1707 DSYSAGQSVEILDGVELGBPAHKKTGTTVPESIHSFIGDGLVKPRALNKKAIQIINRVRD    1766
                                                  | :|::||::|:
Db  2480 H---------DVVKLNP------TSVVESIKRKV-RGLLPHEKI                    2507

Qy  1767 KLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPF                      1808
                             || :| ||||| :||   :||| ||
Db  2508 ------------PLGVEGQVEBLIKQATDPKNLAAMYIGWCPF                      2538
```

RESULT 12
US-10-300-453A-39
; Sequence 39, Application US/10300453A
; Publication No. US20030165934A1
; GENERAL INFORMATION:
; APPLICANT: ELLEDGE, STEPHEN J.
; APPLICANT: CORTEZ, DAVID K.
; APPLICANT: ZOU, LEE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN CHECKPOINT SIGNALING
; FILE REFERENCE: P0233US
; CURRENT APPLICATION NUMBER: US/10/300,453A
; CURRENT FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 60/331,821
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1

```
Db  1264 --LHERLMNGSIDSGASRQDNRDDYRYGVBEKKEVPKVAPTTARPTSELVTVQITKQRL 1321
Qy   528 ------AWGAARRVSKDDWLEWIRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNA  580

Db  1322 NKDAIMPQWKNENLTSKDEWLQWLMKIRIGPLTYGSSPSLRAASSLGDQHPHLARDLFPA 1381
Qy   581 AFVSCWSELNEDQQDELIRSIELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGI  640

Db  1382 AFMSWWTELDSDVQNDI-TSCLLRAISTGIPELIQTILNIAEFMDHSEKGPLPISHD--- 1437
Qy   641 VLLGERAAKCRAYAKALHYKELEFQKGP----------TPAILBSLSINNKLQQPE  687

Db  1438 -VLGRWAEÇTKAFAKACRYKEMSVLKKSGSNQTTFTRKVKLEPNDCQSLITYANKLNVQE 1496
Qy   688 AAAGVLEYAMKHFGELEIQATWYEKLHEWEDALVAYDKQMDTNKDDPELM---------  737

Db  1497 EAAGVVRYAERNEMNFQMRGRWYEKLNEWEKALGAYELEBEKKKSSCPNLQVYDEKDHLMT 1556
Qy   738 ---------LGRMRCLEAIGEWGQLHQQCCEKWT---------LVNDETQA-----KMA  773

Db  1557 PEEAATAEEARNEMERRCLEALGRWDELNSKSV-VWADQRGNRNDSVRDEINKKQLDHKVA 1615
Qy   774 RMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVTLALHQDLFSLAQQCIDKARDLLDA  833

Db  1616 VIAARGAWAVDNWERMADYVSVISENTQDGAMLRAVVAVHNDENTKAMGLIEKVREMIDS 1675
Qy   834 ELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIVEDW  893

Db  1676 ELTAMANESYERAYIPMVSVQQMAELERAIEYKTRPERRPRIALLWSRRLQGCRRNVBQW 1735
Qy   894 QKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDH-PLPTV  952

Db  1736 QRLIMLRGLVLSPQEMHPLRVKFSSMCRKQGKNSMSRAVIRELLSLPANSDLVRAKAPFD 1795
Qy   953 HPQVTYAYMKNMWKSARKIDAFQEMQOHFVQTMQQQAQHAIATEDQQ---HKQELHKLMA 1008

Db  1796 KPLIVLALAKQLYQDDHKDEAIRALEDLANEHWNKRINPIPKATGRELIPPSTKEPARICA 1855
Qy  1009 RCFLKLGEW----------QLNLQ------------GINESTIPK-------VLQYYSAAT 1040

Db  1856 KVLLKLGEWTELKSKTSNNMQVGELSFVRQQVSPQYRTKESRTPETIAFENTINYYQQAT 1915
Qy  1041 EHDRSKYKAWHAWAVMNFEAVIHYK-HQN------QARDEKKKJRHASGANITNATTAAT 1093

Db  1916 QYDPGWHKVWHKLASTHHYAVCRERPHPTTVISPPQCPQQPKKMHIPP---VTRATSPPP 1972
Qy  1094 TAATATTTASTEGSNSESEEAESTENSPTTPSPLQXKVTEDLSKTLLMVTVP-------- 1143

Db  1973 PA----------------QKSPQQPAPPH-SITEPLS-VIIDYPVPPLGSLVGLP 2009
Qy  1144 --------------AVQGFFRSISLSRGNNLQ 1161

Db  2010 PMPAYLASSNSLPPQHHVSPLSNDSPSNSAENKLYLKHAAHAVRCFAKALMCSPGSRIE 2069
Qy  1162 DTLRVLJTLWPDYG--HWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVG-RLIH 1218

Db  2070 DTLRLMQLWFDHGDDKDQDVVFALTETIPDLPVTTWLERAIPQLMARLDCPDDQKSVQLVL 2129
Qy  1219 QLLTDIGRYHPQALIYLPLTVASKSTTTARHNAANKILKRMCEHSNTLVQQAMMVSEELIR 1278

Db  2130 RVLCEIARHRPQAVIYALTVASRSKDVHRSKNAGTVLEKMEYHSKLVREASLVTEELVR 2189
Qy  1279 VAILWHEMWHEEGLEREASRLYPGER-----NVKGMFEVLEPLHAMERG-PQTLKETSFNQA 1333

Db  2190 CAILWHEQWEDALDDASRVYFIRRLQDNNVQAMFDALRNNNDLMQKGAPTTMKEHSFQQT 2249
Qy  1334 YGRDLMEAQEWCRKYMKSGNVXDLIQAWDLYHVFRRISKQLPQLTSLELQVSPKLLMC 1393

Db  2250 YSSDLKEAGRYVQAFESSGNVKDLNQAWEIYCSVFKKLRDQLATLNSLDLVYVSPNIVSA 2309
Qy  1394 RDLELAVPGTYDPNQPIIRICSIAPSLQVITSKQRPRKLTLMGSNGHFVFLLKGHEDLR 1453

Db  2310 KDLELVVPGTYDPSAPIVSIQSFSSKMNVITSKQRPRKMVIRGSNGLDYQFLLKGHEDPR 2369
```

```
Qy  1454 QDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDY 1513
Db  2370 QDERVMQLFGLVNTLLANNSETCRRNLTIQRYSIVALSKDSGLIGWPNCDTLHTLVKEY 2429

Qy  1514 REKK-KILLNIEHRIMLRMAPDYDHLTLMQKVRVFEHAVNNTAGDDLAKLLMLKSPSSEV 1572
Db  2430 REKKAKIPLSIPHKTLQKSLETEHLTLMQKLQLFESALSVTQGELRHVLWLKSPSSEV 2489

Qy  1573 WFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMDRLSGKILHIDFGDCFEVAMTREKFP 1632
Db  2490 WFDRRTNYTRSVACMSMVGYILGLGDRHPSNLMLDRLTGKVVHIDFGDCFEVAMLEKFP 2549

Qy  1633 EKIPFRLTRMLTNAMEVTGLDGNYRITCH-VMEVLREHKDSVMAVLEAFVYDPLLNWRLM 1692
Db  2550 ERVPFRLTRMLINAMEVTGLDGVYNYTAERVLKMLRTNQESLLAVLEAFVYDEVINWRLV 2609

Qy  1693 DTNTKGNXRS-RTRTDSYSAGQSVEILDGVELGGEPAHKKTGTTVPESIHSFIGDGLVKPE 1751
Db  2610 E----GMKKDPKTRKD--TGGR------------------QNMAGAVLPSS----- 2636

Qy  1752 ALNKKAIQIINRVRDKLTGRDFSHDD------TLDVPTQVELLIKQATSHENLCQCYIGW 1805
Db  2637 ---STTDSIMETIKRKLDGTEFVHTDGSTPPEPLQVTEQLAMLTEQATSPLNLCQSYIGW 2693

Qy  1806 CPFW 1809
Db  2694 CPFW 2697
```

RESULT 11
US-10-369-493-3779
; Sequence 3779, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;   APPLICANT: Cao, Yongwei
;   APPLICANT: Hinkle, Gregory J.
;   APPLICANT: Slater, Steven C.
;   APPLICANT: Goldman, Barry S.
;   APPLICANT: Chen, Xianfeng
;   TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;   TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;   FILE REFERENCE: 38-10(52052)B
;   CURRENT APPLICATION NUMBER: US/10/369,493
;   CURRENT FILING DATE: 2003-02-28
;   PRIOR APPLICATION NUMBER: US 60/360,039
;   PRIOR FILING DATE: 2002-02-21
;   NUMBER OF SEQ ID NOS: 47374
;   SEQ ID NO 3779
;   LENGTH: 2539
;   TYPE: PRT
;   ORGANISM: Neurospora crassa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(2539)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3779

Query Match   9.0%; Score 842.5; DB 15; Length 2539;
Best Local Similarity   21.7%; Pred. No. 1.8e-62;
Matches 452; Conservative 281; Mismatches 728; Indels 621; Gaps 85;

```
Qy   22 SNAPRLIR------PYMBPILKALILKLKDPDPDPNPGVINN------------  57
Db  783 ANFPSIIALLLIQDVPDVEEHAMELLCVISPKFKEPNPGELNSDEKRPVSVRDLAEILR 842

Qy   58 --VLATIGELAQVS-GLEMRK--WVDELFIIIMDMLQDS----------SLLAKRQVALW 102
Db  843 VEPLGVLFELFKASGGADERKKAW---YVTFTDIXKPEADNAQIRKGISILATRSSPLT 898

Qy  103 TLGQLVASTGYVVEPY-RKYPTLLEVLLNFLKTEQNQ-GTRREAIRVLGLL--------- 151
Db  899 EKG---AEPPHIIERPLEKYTLGLVSRLSER=TDIDVPIPDRRRCLRAMEMIRVCESSV 955
```

```
Db  2385 LINPSEL---LRKGAITVEBAANM--EAEQQNETRNARAMLVLRRITDKLTGNDIKRF 2437
                    ::  ||   |:  | ||| :::|| ||| |||||
Qy  1777 DTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
              :   || : |:|||||:||| ||||||||||
Db  2438 NEILDVFSQVDKLIQQATSIERLCQHYIGWCPFW 2470


RESULT 10
US-10-369-493-5036
; Sequence 5036, Application US/10369493
; Publication No. US2003023367SA1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT: Hinkle, Gregory J.
;  APPLICANT: Slater, Steven C.
;  APPLICANT: Goldman, Barry S.
;  APPLICANT: Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE: 2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
;  SEQ ID NO 5036
;  LENGTH: 2697
;  TYPE: PRT
;  ORGANISM: Caenorhabditis elegans
US-10-369-493-5036

Query Match             36.3%;  Score 3417;  DB 15;  Length 2697;
Best Local Similarity   38.2%;  Pred. No. 7.9e-286;
Matches  781;  Conservative  336;  Mismatches  603;  Indels  324;  Gaps  45;

Qy     1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDDPNPGVINNVLA 60
Db   743 MQQSGQARLEQHSAXMIAQLAKQSPKEMRPYVGSLMIAMIPKLRNDQKYAE--VTAQVLN 800

Qy    61 TIGELAQVSGLEMRKWVDRLFIIIMDMLQDSSLLAKRQVALIWTLGQLVASTGYVVEPYRK 120
Db   801 AVSEIAVIGGAEIVKNLPLFEKLTHMINDSSSLHKREAALRAIGGICRSTAYVVDPYRD 860

Qy   121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYYKHKVNIGMIDQSRDASAVSLSES 180
Db   861 YPSLLDDLLRILKTVMSNTMRREAIKTLGLLGAIDPYTHKVFTGSV-QSSTAISTALSLP 919

Qy   181 KSSQDSSDYSTSEML--VNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFI 238
Db   920 ISETDSKD-PRQDIIHWFNYEKCTLBEFYPAITIANLMLMWQDED-SQSYAEIAQAIVTI 977

Qy   239 FKSLGLKCVQFLPQVMPTFLNVIRVCDGA------IREFLFQOLGMLVSFVKSHIRPY 290
Db   978 FRSLGDMAPLYTEQVIP--RLIEVCRRATESSNRANLREFFLQQLANFVAIIRKHAAPY 1034

Qy   291 MDEIVTLMRBFWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPEHMLRVFMHDNSPGR 350
Db  1035 MPAIFTTIIADAWKEDISVKMVVIEVLTDMGTAIGNDFSKYTGEHIPYLLTVLQTDKTKER 1094

Qy   351 IVSIKLLAAIQLFGANLDDYLHELLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTD 410
Db  1095 VLTVKVMESIQKLTHCIVCHLHVLPPLILILDFSLKLSIRNTALSTVLHNTQQVDVSA 1154

Qy   411 YASRII----HPIVRTLDQSPELRSTAMDTLSSIVFQLGKKYQIFIPMVNKVL----VRH 462
Db  1155 YAPRMMQSWHHNI---STAEMRDKLLLIIBIIKQLGKFPDIFKRGVDQKLRDYNLDK 1209

Qy   463 RINHQRY------DVLICRIVKG------YTLADEEEDPL 490
Db  1210 SVHYEQYRKLAQRAQMSRDVLISSVFAGSNGNIQYSSTQAGMRGQANNVYANND------1263

Qy   491 IYQHRMLRSGCQCDALAS----------GPVEGTGPMKKLHVSTINLQK- 527
```

```
Db  1343 AKALIHYKEIKFIKEPENSTIESLISINNQLNQTDAAIGILKHAQQHH-SLQLKETWFEKL 1401
Qy   714 HEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMA 773
Db  1402 ERWEDALHAYNEREKAGDTSVSVTLGKMRSLHALAEWQLSQLAARKNVVSKLQTKKLIA 1461

Qy   774 RMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDA 833
Db  1462 PLAAGARGGSGEWDMLDEYISVMKPKSPDKEFFDAILYLHKNDYDNASKHILNARDLLVT 1521

Qy   834 ELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVP--ERREIIRQIWERLQGCQRIV 890
Db  1522 EISALINESYNRAYSVIVRTQITTEFEEIIKYKQLPPNSEKKLHYQNLWTKRLLGCQKNV 1581

Qy   891 EDWQKILMVRSLVVSPHEDMRLWLKYASLCCGKSGRIALAHKTLVLL--GVDPSRQLDHP 948
Db  1582 DLWQRVLRIRSLVIKPKQDLQIWIKFANLCRKSGRMRLANKALNMLLEGGTTIVYQLRSK 1641

Qy   949 LPTVHPQVTYAYMKNMWKSARKIDAFQEMQHFV--------QTMQQQAQHAIATEDQ 997
Db  1642 PP---PPVVVAQLKYIWATGAYKEALNLHLIGFTSRLAHDLGLDFNNMIAQSVKLSSASTA 1698

Qy   998 QHKQELHKLMARCFLKLGBWQLNLQGINESTIP-KVLQYYSAATEHDRSWYKAWHAWAVM 1056
Db  1699 PYVEEYTKLLARCFLKQGEWRIATQPWRNTNPDAILGSYILLATHFDKNWYKAWHNWALA 1758

Qy  1057 NFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAASTEGSNSESBAEST 1116
Db  1759 NFEVI------SMVQBEFYKLN-GGKNDDDDDTAVN-----NDNVRIDGSILGSGSGLTI 1804

Qy  1117 ENSPTPSPLQKKVTEDLSKTLLMYTVPAVCGFPRSISLSRGNNLQDTLRVLTIWFDYGHW 1176
Db  1805 NGNYPLELIQR--------HVVPAIKGFFHSISLLETSCLQDTTLRLSTLLFNFGGI 1853

Qy  1177 PDVNEALVEGVKAIQDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPL 1236
Db  1854 KEVSQAMYEGFNLMKLENWLEVLPQLISRIHQPDPTVSNSLLSLLSDLGKAHPQALVYPL 1913

Qy  1237 TVASKSTTTARHNAANKILKNMCEHSNTLVQQAMMVSPELIRVALLWHHMWHEGLEEASR 1296
Db  1914 TVAIKSESVSRQKAALSIIEKIRIHSPVLVNQAELVSHELIRVAVLWHELWYEGLEDASR 1973

Qy  1297 LVFGERNVKGMFEVLEPLHAMMERGPCTLKETSPNQAYGRDLMEAQEWCRKYMKSGNVKD 1356
Db  1974 QFFVEHNIEKMYSTLEPLHKHLGNEPQTLSEVSFQKSFGRDLNADAYEWLNNYCKSKDINN 2033

Qy  1357 LTQAWDLYYHVFRRISKQLPQLTSLELQYYVSPKLLMCRDLBLAVPGTYDPNQPIIRIQSI 1416
Db  2034 LNQAWDIYYNVFRKITRQLPQLQTLDLQHVSPQLLATHDLBLAVPGTYPFGKPTIRIAKF 2093

Qy  1417 APSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSL 1476
Db  2094 EPLFSVISSKQRPRKFSIKGSDGKDYKYVLKGHEDIRQDSLVMQLFGLVNTLLKNDSECF 2153

Qy  1477 RKNLSIQRYAVLPLSTNSGLIGWVPHCDTLHALLIRDYREKKKILLNIEHRIMLRMAPDYD 1536
Db  2154 KRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHDAKKIPLNIEHVMLQMAPDYE 2213

Qy  1537 HLTLMQKVEVFEHAVNNTAGDDLAKLLJLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGL 1596
Db  2214 NLTLLQKIBVFTYALDNTRKGQDLYKLIIWLKSRSSETWLERRTTYTRSLAVMSMTGYILGL 2273

Qy  1597 GDRHPSNLMLDRLSGKILHIDFGDCFFVAMTREKFPEKIPFRLTRVLTNAMEVTGLDGNY 1656
Db  2274 GDRHPSNLMLDRITGKVIHIDFGDCFEAAILREKYPEKVPFRLTRMLTYAMEVSGIEGSF 2333

Qy  1657 RITCHTMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKSRRTRTDSYSAGQSVE 1716
Db  2334 RITCENVMRVLRDNKESLMAILEAFALDPLLHWGFDLPPFQKLTEQT------GIPLP 2384

Qy  1717 ILDGVELGBPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHD 1776
```

```
;  APPLICANT: Start, William G.
;  APPLICANT: Tennesen, Dan
;  APPLICANT: Vidya, K.R.
;  APPLICANT: Wang, Haiyun
;  APPLICANT: Xin, Zhanguo
;  APPLICANT: Xu, Nanfei
;  APPLICANT: Yang, Chunzhi
;  APPLICANT: Zeng, Xiaoping
;  APPLICANT: Zhang, Qiang
;  APPLICANT: Zhao, Yajuan
;  APPLICANT: Zhou, Li
;  TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
;  FILE REFERENCE: 38-15(52796)B
;  CURRENT APPLICATION NUMBER: US/10/310,154
;  CURRENT FILING DATE:  2002-12-04
;  PRIOR APPLICATION NUMBER: 60/337,358
;  PRIOR FILING DATE: 2001-12-04
;  NUMBER OF SEQ ID NOS: 736
;  SEQ ID NO 579
;  LENGTH: 2470
;  TYPE: PRT
;  ORGANISM: Saccharomyces cerevisiae
US-10-310-154-579

Query Match             41.6%;  Score 3913;  DB 15;  Length 2470;
Best Local Similarity   44.0%;  Pred. No. 0;
Matches  806;  Conservative 362;  Mismatches 561;  Indels 104;  Gaps 25;
```

```
Db    2140 QLFGLVNTLLQNDABCFRRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHREAKKI 2199          Db    942 QTLGLKCVSFLDQIIPTILDVMRTCSQSLLEFYFQQLCSLIIIVRQHIRPHVDSIFQAIK 1001
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                         ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1520 LLNIEHRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDIAKLLWLKSPSSEVWFDRRTN 1579          Qy    300 EFWVMNTSIQSTILLILEQIVVALGGEFKLYLPQLIPEMLRVFMEHDNSPGRIVSIKLLAA 359
            ::|:||||||||||||||||||||||||||||||||:::::||:|||||::::|:||                         :|::||||||||||||||||||||||||:|::|::::|:||:|::|:::||
Db    2200 PINIEHWVMLQMAPDYDNLITLLQKVEVFTYALNNTBGQDLYKVTLWLKSRSSETWIERRTT 2259          Db   1002 DFSSV-AKLQITLVSVIEAISKALEGEFKRIVPLTLTLFLVILENDKSSDKVLSRRVLRL 1060

Qy    1580 YTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPPEKIPPRL 1639          Qy    360 IQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPI 419
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                         ::|||||||||||||||||||||||||||||||||||||||||||||||::|||
Db    2260 YTRSLAVMSMTGYILGLGDRHPSNIMLDRITGVIHHIDFGDCFEAAILREKFPEKVPFRL 2319          Db   1061 LESFGPNLEGYSHLITPKIVQMAEFTSGNL--QRSAIITIGKLAKDVDLFEMSSRIVHSL 1118

Qy    1640 TRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGN 1699          Qy    420 VRTLDQ--SPELRSTAMDTLSLLVFQLGKKYQIFIPMVNKVLVRRHIRHQRYDVLICRIV 477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                         ::|||||||||||||||||||::|||||::||||||||||||::||||||||
Db    2320 TRMLTYAMEVSGIEGSFRITCENVMKVLRDNKGSLMAILEAFAFDPLINWG-FDLPTK-- 2376          Db   1119 LRVLSSTTSDELSKVIMNTLSLLIQMGTSFAIFIPVINEVLMKKHIQHTIYDDJLTNRIL 1178

Qy    1700 KRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVP-ESIHSFIGDGLVKPEAL----- 1753          Qy    478 KGYTLADEEDPLIYQHRMLRSGQGDALASGPVETGP--MKKLHVSTINLQKAWGAARRV 535
            ::|||||||::|||||||||||:::||||||                                   ::|||||||::|||||::||||||||||||
Db    2377 ------------------KIBEETGIQLPVMNANELLSNGAITEEEVQRVEN 2410          Db    478 KGYTLADEEDPLIYQHRMLRSGQGDALASGPVETGP--MKKLHVSTINLQKAWGAARRV 535

Qy    1754 -----NKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWC 1806          Db   1179 NNDVLPT---------KILEANTIDYKPAEQMEAADAGVAKLPINQSVLKSAWNSSQQR 1228
            ::|:|||||::||::||||||||||::||||||
Db    2411 EHKNAIRNARAMILVLKRITDKITGNDIRRFNDLLDVPEQVDKLIIQQATSVENLCQHYIGWC 2470          Qy    536 SKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQD 595

Qy    1807 PFW 1809          Db   1229 TKEDWQEWSKRLSIQLLKESPSHALRACSNLASMSYYPLAKELFNTAFACVWTELYSQYQE 1288
            |||
Db    2471 PFW 2473          Qy    596 ELIRSIELALTSQ-DIAEVTQTLLNLAEFMEHSDKG-PLPLRDDNGIVLLGERAAKCRAY 653
                                                                                          ::|:|||||||||::||||||||||||
                                                                                   Db   1289 DLIGSLICIALSSPLNPPEIHQTLLNLVEFMEHDDKALPIPTQS-----LGEYAERCHAY 1342

RESULT 8                                                                         Qy    654 AKALHYKELEFQKGPTPAILESLISINNKLQQPBAAAGVLEYAMKHFGELEIQATWYEKL 713
US-10-369-493-22191                                                                          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
; Sequence 22191, Application US/10369493                                        Db   1343 AKALHYKEIKFIKEPENSTIRSLLSINNQLNQTDAAIGILKHAQQHS-SLQLKETWFEKL 1401
; Publication No. US20030233675A1
; GENERAL INFORMATION:                                                           Qy    714 HSWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKWA 773
;  APPLICANT: Cao, Yongwei                                                                   ::|||||||||::|||||::||||||||||||
;  APPLICANT: Hinkle, Gregory J.                                                 Db   1402 ERWEDALHAYNEREKAGDTSVSTVILGKVMRSLHALGEWEQLSQLAARKWKVSKLQTKKLIA 1461
;  APPLICANT: Slater, Steven C.
;  APPLICANT: Goldman, Barry S.                                                  Qy    774 RMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDA 833
;  APPLICANT: Chen, Xianfeng                                                                 ::|:||||:|||:||::||||||
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  Db   1462 PLAAGAAWGLGEWDMLEQYISVMKPKSPDKEFFDAILYLHKNDYDNASKHILNARDLLVT 1521
;  TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(152052)B                                                Qy    834 ELTAMAGESYSRAYGAMVSCHMLSELEEVIQVKLVP---ERRELIRQIWWERLQGCQRIV 890
;  CURRENT APPLICATION NUMBER: US/10/369,493                                                 :::||:|||::|||::||||||
;  CURRENT FILING DATE: 2003-02-28                                               Db   1522 ETSALLNESYNRAYSVVRTQIITBFEEIIKYKQLPPNSSKKLHYQNLWTKRLLGCQKNV 1581
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21                                                 Qy    891 EDWQKILMVRSLIVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLIL--GVDPSRQLDHP 948
;  NUMBER OF SEQ ID NOS: 47374                                                               ::|||::||||::|||||||||
; SEQ ID NO 22191                                                               Db   1582 DLWQRVLRVRSLIVLKPKQDLQIWIKFANLCRKSGRMRIANKALNMLLEGGNDPS------ 1635
;  LENGTH: 2470
;  TYPE: PRT                                                                     Qy    949 LPTVH---PQVTVAYMQNWKSARKIDAFQHMQHFV-----------QTMQQQAQHAIAT 994
;  ORGANISM: Saccharomyces cerevisiae                                                        :||||
US-10-369-493-22191                                                             Db   1636 LPNTFKAPPPVVYAQLKYIWATGAVKEALNHLIGFTSRLAHDLGLDPNNMIAQSVFLSSA 1695

Query Match     42.1%;  Score 3962;  DB 15;  Length 2470;                       Qy    995 EDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIP-KVLQYYSAAETEHDRSWYKAWHAW 1053
Best Local Similarity   44.4%;  Pred. No. 0;                                                 ::|||||||||||::||||||
Matches  815;   Conservative  360; Mismatches  551; Indels  110;  Gaps    26;   Db   1696 STAPYVBEYTKLLARCFIKQGEWRIATQPNWRTNPDAILGSYLLATHFDKNWYKAWHNM 1755

Qy       1 LEHSGIGRKIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDDPDNPGVINNVLA 60           Qy   1054 AVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATTAATTASTEGSNSESEA 1113
            |::|:|||||||||||||:|||::|:||||||                                         ::||||||||||||::|||||||
Db     718 LKFSTSSREKBETASLLCTLIRSSKDVAKPYIEPLLNVLLPKFQ----DTSSTVASTALR 773          Db   1756 ALANFEVI-------SMVQEETKLN--GGKNDDDDTTAVN-----NDNVRIDGSILGGGS 1801

Qy      61 TIGELAQVSGLEMRKKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120          Qy   1114 ESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLLTLWFDY 1173
            |||||||||||||||||||:||||||||||::|||||||                                 :::||||||||||::|||||||
Db     774 TIGELSVVGGBDMKIYLKDLFPLIIKTFQDQSNSFKREAALKALGQLAASSGYVIDPLLD 833          Db   1802 LTINGNRYPLELIQR-------HVVPAIKGFFSISLLETSCLQDTLRLITLLTLLPNF 1850

Qy     121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMDQSRDASAVSLSES 180          Qy   1174 GHWPDVNEALVEGVKAIQIDTWLQVIPQLLIARIDTPRPLVGRLIHQLLTDIGRYHPQALI 1233
            |||||||||||:||||::|||||||||                                   :|:||||||||||||||::|||||||
Db     834 YPELLGILVNILKTENSQNIRRQTVTLIGILGAIDPYRQK---------EREVISTTDI 883          Db   1851 GGIKEVSQAMYEGFNLMKIENWLEYLLPQLISRIHQPDPTVSNSLSLSLSLDLGKAHPQALV 1910

Qy     181 KSSQDSSDYSTSEMLVNMGNLPL-DEFYPAVSMVALMRIFRDQSLSHHTMVVQAITPIF 239          Qy   1234 YPLTVASKSTTTARHNAANKIIKNMCEHSNTLVQQAMMVSEELIRVAILWEEMWHEGLEE 1293
            ::|||||||||||::|:||||||                                       :::||||||||||||::|||||||
Db     884 STEQNAPPIDIA--ILMQGMSPSNDEFYTTVVIHCLLKILKDPSLSSYHTAVIQAIMHIF 941          Db   1911 YPLTVAIKSESVSRQKAALSIIEKIRIHSPVLVNQAELVSHELIRVAVLWEELWYBGLED 1970

Qy     240 KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDBIVTLMR 299          Qy   1294 ASRLYFGERNVKGMFEVLEPLHAMVERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGN 1353
            |: ||:||| :: |: ||||:|                                       ::|||||||||||::|||||||
                                                                         Db   1971 ASRQFFVEHNIEKMFSTLEPLHKHLGNEPQTLSEVSFQKSFGRDLNDAYEWLNNYKKSKD 2030
```

```
Db  2133  LGLGDRHFSNLMMDRYSGKIIHIDFGDCFEVAMHREKPPEKIPFRLTRMLINAMEVSGIQ  2192
Qy  1654  GNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQ  1713
Db  2193  GTYKITCELVMRVLRSNTESLMAVLEAFVYDPLINWRLMTKSSFGASTTLRPTSS-----  2247
Qy  1714  SVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQLINRVRDKLTGRDF  1773
Db  2248  SVE-------EKGRSYTHRARHADYAALSETNGVNAEGLNERSIQVLKRVSNKLTGKDF   2299
Qy  1774  SHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW  1809
Db  2300  DLKEQLPVKAQVEKLIQQATAPENLCRCYVGWCSFW  2335
```

```
Query Match          43.6%;  Score 4105;  DB 15;  Length 2473;
Best Local Similarity 45.5%;  Pred. No. 0;  Indels 170;  Gaps 28;
Matches  847;  Conservative 333;  Mismatches 513;  Indels 170;  Gaps 28;

Qy     1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPIIKALLILKDPDPDNPGVINNVLA   60
Db   727  LKFSNMPKKKERSATLLCTLINSSDEVAKPVIDPILDVIILPKCQ----DASSAVASTALK  782

Qy    61  TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK  120
Db   783  VLGELSVVGGKEMTRYLKELMPLIINTFQDQSNSFKRDAALTTLGQLAASSGYVVGPLLD  842

Qy   121  YPTLEVLNFLKTEQNGQTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES   180
Db   843  YPELGGTILINILKTENNPHIRRGTVRLLIGILGALDPYKHR------------EIEVTSNS  890

Qy   181  KSSQDSDSDYSTSEMLVNMGNLPL-DEFYPYAVSMVALMRIFRDQSLSHHHTMVVQAITFIF  239
Db   891  KSSVRQNAPSIDIALLMQGVSPSNDEYPTVVIHNLMKILNDPSLSTHHTAAIQAIMHIF   950

Qy   240  KSLGLKCVQPLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR  299
Db   951  QNIGLRCVSFLDQIIPGILVMRSCPPSQLDFYFQQLGSLISIVKQHIRPHEVEKIYGVIR  1010

Qy   300  EFWVMNTSIQSTIILLIEQIVALGGEFKLYLPLQIPHMLRVFMHDNSPGRIVSIKLLiAA  359
Db  1011  EFFPI-IKLQITIISVIBSISKALRGEFKRFVPETLTFFLDILENDQSNKRIVPIRILKS  1069

Qy   360  IQLFGANLDDYLHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPI   419
Db  1070  LVTFGPNLEDYSHLIMPIVVRMTRYSAGSL--KKISIITLGRLAKNINLSEMSSRIVQAL  1127

Qy   420  VRTLDQSP-ELRSTAMDILSSIVFQLGKKYQIPMVNKVLVRHRINHQRYDVLICRIVK   478
```

```
Db  1128  VRIINNGDRELTKAIMNTLSLLLLQLGTDFVVFPVINKALLRNRIQHSVYDQLVNKLLN  1187
Qy   479  GYTIA-----DEERDPLIYQHFMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAA   532
Db  1188  NECLPTNIIFDKENEVP------ERRNYEDEM------QVTKLPVNQNILKRNANYCS   1232
Qy   533  RRVSKDDWLEWLRRLSLELIKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNED  592
Db  1233  QQKTKEDWQEWIRRLSIQLLKESPSACLRSCSSLVSVYPLARELFNAGFSSCWVELQTS   1292
Qy   593  QQDELIRSTELAL-TSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCR  651
Db  1293  YQEDLIQALCKALSSSENPPEIYQMLLNLVEFMEHDDK-PLPI----PIHTLGKYAQKCH  1347
Qy   652  AYAKALHYKELHFOKGPTPAILESLLSISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYE  711
Db  1348  AFAKALHYKEVEFLEEPKNSTIBALISINNQHQTDSAIGILKHAQQH-NELQLKETWYE  1406
Qy   712  KLHBWEDALVAYDKKMDTNKDDPELMLGRMCLEALGEWGQLHQQCCEKWTLVNDETQAK  771
Db  1407  KLQRWEDALAAYNEKEAAGBDSVEVMMGKLRSLYALGEWEELSKALASEKWGTAKPEVKKA  1466
Qy   772  MARMAAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVLALHQDLFSLAAQCIDKARDLL  831
Db  1467  MAPL-AAAAWGLEQWDEIACYTSVMKSQSPDKEFYDAILCLHRNNFKKAEVHIFNARDLL  1525
Qy   832  DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVP--ERREIIRQIWERLQCQR   888
Db  1526  VTELSALVNESYNRAYNVVVRAQIAELEBIIKYKKLPQNSDKRLTMRETWNTRLLGCQK  1585
Qy   889  IVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLGVDPSRQLDHP  948
Db  1586  NIDVWQRILRVSLVIKPKEDAQVRIKFPANLCRKSGRWALAKKVLNTLL--EEYDDPDHP  1643
Qy   949  -LPTVHPQVTYAVMKNMWKSARKIDAFQHMQHF-------VQTMQQQAQHA   991
Db  1644  NTAKASPPVVYAQLKYLWATGLQDEALKQLINFTSRWAHDLGLDPNNMIAQSVPQQSKRV  1703
Qy   992  IATEDQQHKQELHKLMARCRLKLIGEWQLNLQGINESTIPK------VLQYYSAAIEHD  1043
Db  1704  ------PRHVEDYTKLLLARCFLKQCEWRVCLQ----PKWRLSNPDSILGSYLLATHFD  1751
Qy  1044  RSWYKAWHAWAWMNFEAVLHYKHQNQARDEKKGLRHASGANITNACTAATTAATATTTAS  1103
Db  1752  NTWYKAWHNWALANFEVI------------------SMLTSVSKKK  1779
Qy  1104  TEGSNSESEAESTE---NSPTPSPLQKKVTEDLSKTLLMYTTVPAVQGFFRSILSRGNNL  1160
Db  1780  QEGSDASSVTDINEFDNGMIGVNTFDAKEVHYSSNLIHRHVIPAIKGFFHSISLSESSSL  1839
Qy  1161  QDTLRVLIMFDYGEWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQL  1220
Db  1840  QDALRLLTLWFTFGGIPEATOAMHEGENLIQIGTWLEVLPQLISRIHQPNQIVSRSLLSL  1899
Qy  1221  LTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSELIRVA   1280
Db  1900  LSDLGKYAHPQALVYPLMVAIKSELSRQKAALSIIEKMRIHSPVLVDQAELIVSHELIRMA  1959
Qy  1281  ILWHEMWHEGLEEASRLYFGBRNVKGMPEVLEPLHAMMERGPQTLKETSFNQAYGRDLME  1340
Db  1960  VLWHEQWTYGLDDASRQFFGEHNTEKMFAALEPLYEMLARQPTLREISPQNSFGRDLND  2019
Qy  1341  AQEWCRKYMKSGNVKDLTQAWDLYYEVRRISKQLPQLTSLELQYVSPKLLMCRDLELAV  1400
Db  2020  AYEWLMNYKKSKDVSNLNQAWDIYYNVFRKIGKQLPQLTLELQHVSPKLLSAHDLELAV  2079
Qy  1401  PGT-YDPNQPIIRIQSIAPSLQVTSKQRPRKLITLMGSNGHEFVFLLKGHEDLRQDERVM  1459
Db  2080  PGTRASGGKPIVKISKFEPVFSVISSKQRPRKFCLKGSDGKDYKYVLKGHEDIRQDSLVM  2139
Qy  1460  QLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVFHCDTLHALIRDYREKKKI  1519
```

```
Qy   1734  TVPESIHSFIGDGLV--------KPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVFTQ 1784
            :: |:   ||                  :||: |:::|:: ||.|||:||
Db   2139  TNSSAIYP-TGSSMVNGHGNNQGEQQEFVQNARALEVLDRVQQKLTGRDFKPHEELNVTDQ 2197

Qy   1785  VELLIKQATSHENLCQCYIGWCPFW 1809
            ||:||:||||:||||::  ||||
Db   2198  VNKLIIEATKLENLCQHYIGWCSFW 2222
```

RESULT 6
US-10-369-493-2561
; Sequence 2561, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2561
; LENGTH: 2335
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2561

  Query Match   44.1%; Score 4147; DB 15; Length 2335;
  Best Local Similarity 45.9%; Pred. No. 0;
  Matches 843; Conservative 337; Mismatches 516; Indels 140; Gaps 24;

```
Qy     1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKDPDPDPNPGVINNVLA 60
          :|:|| ||:|| ||||| :||:  ||  |:|  :|:::
Db   613  MEYSGNSRQKEPSAQLLKLLVSKARTLIKPYIQSIIHVILPKA---ADTSPGVSSAIIS 668

Qy    61  TIGELAQVSGLEMRKWVDELFI-IIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYR 119
          :|| |  ||||:| |||| ||| |  :|  |::| : |  | ||:| |: |   |||
Db   669  ALGELASVEGEDMPVDVRGSFMKLILVNLQDQSSTLKRLASLKCLRKLCGRSGYVIQPYL 728

Qy   120  KYPILLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKYNIGMIDGSRDASAVSLSE 179
          ||| ||:||| |:|:||   ||::|  :||: ||:    |  |:|| |  | ||:|: |
Db   729  DYPPLLGALIGILQSEQPTPIRRRVLRTLGVLGALDPYTY---------LTTEEVSD 776

Qy   180  SKSQDSSDDYSTSEMLVNMGNLPLDEYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIF 239
          :|:|| :     ||:|:  ||  ::    | ::|::|   ::| |:  ||:|| ||  |
Db   777  DLQSSHNNAHGVPQ--ISAAQYPSLENYAMVAVVTLIGILKDSSLSMHHSSVVQAVMHIC 834

Qy   240  KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR 299
          || |:| ||| ||  |:|  |:::|  |  |  :|||   |::  |: |:|||  |  |:
Db   835  SQMGSKSTVFLPQVVPTFLQVMQSLSASSABFYFQQLTTSIIGPNIRDYVSDIFNLSK 894

Qy   300  EFWYMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPEMLRVFMHINSPGRIVSIKLLAA 359
          ::| |||||::|||| | :| || ::||  |  |:  |::|:|:|||| :| :|:||
Db   895  VFWESTTSLLLVILELIDAIAIALQDEFKFYLPQILSCMLKAFSLDNTSSRSVSYKVLQS 954

Qy   360  IQLPFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPI 419
          :|:|| |:|  | |:|| |::||  ||   ||:||:::|  ||:|:   ||  | ||:||:
Db   955  FVIFGSNIBEYMHLVLPVIIRSFRDTIPIGFRKSALKCIAQLFQSVNFSDHASRIIHPL 1014

Qy   420  VRTLDQS-PELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVPHRINHQRYDVLICRIVK 478
          || ||:|  :|| |  || | |:  ||: |:|||:|||:||:   :|: ||||:|  :|
Db  1015  VRMLGKSNGDLRAVIMDTLCAIVSQLGVDYSIFIPMVNKVLVSHKISHPYELLVSRLLK 1074

Qy   479  GYTLADEEEDPLIYQHRMLRSGGQGDALASGPVETGP-MKKLLHVSTINLQKAWGAARRVS 536
           |  |::   :| :   |: |  :  |:::| |  |  |||||| |:|  | | |  |:|
Db  1075  GEPL---PKDVVVKEFKPRPS--------TKPFSTQDEVLTKLPVDCQASLKAAWESSQKLT 1124
```
```
Qy   537  KDDWLEWLRRISLELLKDSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDE 596
          ::|| |  |||:|||:|:|||||:|:||||| || |  ||:: |  :|||:|::  ::
Db  1125  RDDWQDWIRRISIELLKSPSSALRSCSTIAGIYHPLARDIFNVSFLSCWDELTESNKKN 1184

Qy   597  LRRSIELALTSQDIA-EVTQTLLNLAEFMEHSDKG-PLPLRDDNGIVLLGERAAKCRAYA 654
          |::|::|||  |||: ::|:|||:||:||:  : : | |: ::  :|:|: ||: ||
Db  1185  LVKSIELAMNAPNISVEIIQTLLNLAEYMEREDHTLPIPIK-----VISAHASKCNVYA 1238

Qy   655  KALHYKELEF----QKGPTPAILESLISINNKLQQPEAAAGVLEYAMCHFGELEIQATWY 710
          ||||: |:|:    |:  | |  || :| :||:|:: |  ||   : |: |  ::| |:
Db  1239  KALHYTELQFYQFTKEEVSISTIESLITINNHLQQSDAAVGMLQYTKEH-KQFSLKETWY 1297

Qy   711  EKLHBWEDALVAYDKKMDTNKDDPEIMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQA 770
          ||||: || |||:|::   ::| | | : : |  | | || |: |::  | |  :| |
Db  1298  EKLHRWDDALAAYEHREREGDSSFEINIGKLRCYALGDWDHLSELAQKAWTSEQEHRE 1357

Qy   771  KMARMAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVILAHQDLFSLAQQCIDKARDL 830
          |   | ||||||:|||| | :| :|  :|  | : |  :|||:|::|:||  | :| |
Db  1358  AIAPLAAAAAWGLGQWNLISEYVSAMDRDPQDKEPFSAISAVHLGQYNKAYGHIERHRDI 1417

Qy   831  LDAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYK--LVPERREIIRQIKWWERLQGCQ 887
          | ::| |:|| :  || | :::|::|::| :::||    :|:|:: |:   |::|| ||
Db  1418  LVNDLSSIIGESYNRAYGIMVKSQMLSELEEIIDYKKNMQYENNLDSLKKTWRKRLEGCQ 1477

Qy   888  RIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVILLGVDPSRQLDH 947
          :||::||::|::|||  :| : | ||||:|:|::| :|:  |:|:|::|||||:||: : |
Db  1478  KNVDVWHNTLRFRALVLSPQDSPEMWIKLADLCRRSDRLKLSNQCLTYLMGRDPSNA--Y 1535

Qy   948  PLPT---VHPQVTYAYMGNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQH----- 999
          || |    ||  :  |  :|: ::  ::  : :  :: :|: ::
Db  1536  PLDSLKLLNPHVVYTYLKYLWATDQKNIAVSELEEF--FSYLSSKHGYKMGDSSKLVDIL 1593

Qy  1000  ------KQELHKILMARCFLKLGEWQLNLQ--GINESTIPKVLQYYSAATEHDRSWYXAWHAW 1053
                ::| : :|: :: :|:||::||:|  |:||| :| :|::||:  |  :|  |:|: |
Db  1594  ASSSVSSEERSFLARCFHKLGKWKKSLQDSVNQESVRDILNCYFYATLPDKSWYKAWHSW 1653

Qy  1054  AVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAAITPATTTASTGGSNSESEA 1113
           |:|:
Db  1654  ALANFEVVGYY-------------------------------- 1664

Qy  1114  ESTENSPTPSPLQKKVTEDLSKTLLMYTTPAVQGFFRSISLSRGNNLQDTLRVLTLWFDY 1173
          |: ::  ||::|::  |    |:| |   : | | |:|:|:  :|: |::||:|  | |
Db  1665  EQTEHG----------VTQDMYE--QYIVPAIKGFFHSSVLNQKNSLQDILRLLNLWFKF 1712

Qy  1174  GHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALI 1233
          ||  :||:|: |||::| |||||:| ||:| | |::  :::::|||:| ||||:|:|||:
Db  1713  GEHSDVAAAIVEGFSNVPMDTWLEVIPQLIARIHTSSSSVRASVHQLLSDIGRVHPQALV 1772

Qy  1234  YPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSELIRVALIWHEMWHBGLEE 1293
          || | ||:||  :|: | :: |: :|: | :|::|::|||::|| |:|| ||:||||:
Db  1773  YSLTVSSKSTNPQQKHSAKSIMDSMLSHSDTLVRQALIVSQELIRVAILWHBLWYEGLEE 1832

Qy  1294  ASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKTSFNQAYGRDLMEAQEWCRKYMKSGN 1353
          ||| :  :|::| | ||| || :  :| ||:|:   |:||:|:|:: | |||:|: ||
Db  1833  ASQAYFSDHDISLMIDIVKPLHETLEKGPSTLSEISFAQTFGYDLRKARSYWQKFLQDGD 1892

Qy  1354  VKDLTQAWDIYHVFRRISKQLPQLTSLELQVYSPKLLMCRDLELAVPGTYDPNQPIIRI 1413
          : ||| || :|| ||: |||| :|| :||:|||| ||:|  ||||| |||:||:||:|:
Db  1893  PTELNQSWDLYYQVFRRIQKQLPRIKGHELQYVSPKLLDACDLELAVPGTYGHNKPVIRI 1952

Qy  1414  QSIAPSLQVITFSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDP 1473
          :|: | ||:|: ||:|  |:| ||::|:|: :|:|| :|:| | |: |: ||: ||  |
Db  1953  SHFHHTFEVISSKQRPRRLTIHGSDGKDYQVLKGHEDLRQDERVMQLFGLCNTLLTTDS 2012

Qy  1474  TSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRMAP 1533
          ||: |:|  || :|:| |:|||| || |::|:|::| |: :  | : ||::|| |::| |
Db  2013  ETFKRRLNIERTTVIPLSPNSGLLGWVPHSDTLHFLIKEFRSKRNILANLEHRMMLQMAP 2072

Qy  1534  DYDHLTLMQKVEFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSWVGYI 1593
          |  | |:||:|:| |  ::| || | |::|:|::||: :| ||| :|:: :|:|:|||:
Db  2073  DCDSLTLLQKLEVFFYVMANTDGYDLYHVLWLKSRSSAWLDRRTSYTQSLAVMSWVGYI 2132

Qy  1594  LGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPEKIPFRPLTRMLTNAMEVTGLD 1653
```

```
; CURRENT FILING DATE:  2003-02-28
; PRIOR APPLICATION NUMBER:  US 60/360,039
; PRIOR FILING DATE:  2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3923
; LENGTH: 2222
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2222)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3923

Query Match             44.1%;  Score 4155;  DB 15;  Length 2222;
Best Local Similarity   44.3%;  Pred. No. 0;
Matches  888;  Conservative  317;  Mismatches  520;  Indels  280;  Gaps  34;

Qy     1 LEHSGIGRIKEQSARMLGHIVSNAPRLIRPYMEPILKALILKKPDPDPNPGVINNVLA 60
           ||  |  ||| ::: || : | : ::||  : : :|:||:       :|     ||:
Db   302 LEYTDVARSKEESSKLLSLLTQHAQDLVKPYVNSITEVLLPKAR---DPTPSVAATVLQ 357

Qy    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
           |||| ::: | |: ||:|| :| |::: ||| | |: ||:| |||| |: |:| |::|
Db   358 AIGELCTVGGESMLAYKDTIMPIIIDALQDQSAPIKREAALHTLGQLASNAGTVIKPYLE 417

Qy   121 YPTLLEVLLNFLJKTEQNGT-RREAIRVLGLLGALDPYKHKV----- 161
           ||:|||:| :: ::  : || |:|:| ||:||||||||||:|
Db   418 YPQLLEILQSIIRGEPQHGTLRQETIKLMGILGALDPYKHQVXADTLRSIPDEIKHLRHLT 477

Qy   162 --------NIGMIDQ-SRDASAVSLSESKSSQDSSDYSTSEML--- 195
Db   478 DHLYEMLQASPAKDSMASQVANVASARSGNPSNKPDSLKTALTNLMQVEERTDDNKRNN 537

Qy   196 --------VNMGNL-PLDE-FYPAVSMVALMRIPRDQSLSHHHTMVVQAITFIFKSLGL 244
           :  | : | || || |: || || || ||||:: :| |:| :|
Db   538 EAAQLTDVSLMMGGLTPSNEDYYPTVVINALLQILKDQSLVQWHGNVVDAIMSIFITIGL 597

Qy   245 KCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDRIVTLMREFWVM 304
           ||||||  | | || :||   ::| |: :: |:||| :|||||||:| :|::|:
Db   598 KCVQFLDRVVPAFISVIRASSNARLEFYFNHLSRLVSIVRQHIRVLVKDIIDVLQEYVHT 657

Qy   305 NTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFWHDNSPGRIVSIKLLAAIQLFG 364
           :|::|||| ||||  :|| | ||| | :||:|| |:|:  |||: | || |:  |:| |
Db   658 TLALQSTILGLIESISRSLEBGEFKIYLAKLLPLMLGVLEKDVSTKRQPSEKVFHAFLVFG 717

Qy   365 ANLDDYLHLLPPTVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIVRTL- 423
           :: :||:|| | ::|||| ::| ||||:|||| :| |:| || | ||| |||:||:|
Db   718 SSAEBYMHLIIPVIVRLFDSHSQPLFLRKSAIETIGKLSSMVNLNDYASKIIHPLTRVLA 777

Qy   424 DQSFELRSTAMDTLSLSLVFQLGKKYQIPIPMVNKVLVRHRINHQRYDVLICRIVKG---- 479
           :|  |||  | ||| | || ||: |:|  :   :  |::| ::| |::|: :::|
Db   778 SGEFSLRVAALDTLCAMLQLGRDYLHEHTVDKAISMYAIQHSNYEKKAIEKLKKGEALP 837

Qy   480 YTLADEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDD 539
           | || : | |: | :   |   ::|  :|||   | :|| | ::||::: :  ::|:|||
Db   838 QNLAPRFED--------ISMEGFA-----AENNPPKKLTLNPVHLKQVWETKGKSTKDD 883

Qy   540 WLEWLRRLSLELLKDSSSSPSLRSCWALACAYNPMARDIFNAAFVSCWSELNEDQQ---- 594
           | || ::: ::| |  | :|||:|| |:| | ||:|| ||| || |:|||:|: |:
Db   884 WHEWPRKFSTTLLTESPNHSLRACASLASNYQPLARELFNSAFVSCWSEIYEQFQVCQRG 943

Qy   595 ----------DELIRSIELALTSQDI-AEVTQTLLNLAEFMEHSDKG-PLPLRDDN 638
                      :|:| : | |:| :: :|||:|||||||||:| || |:|| ::
Db   944 IXNDPSRILTACLTGDLITNIENTIKSENVPPLLGQLLNLAEFMEHDDKALPIDIR--- 1000

Qy   639 GIVLLGERAAKCRAYAKALHYKELEFQKGPTPAILRSLLISINNKLQQPEAAAGVLEYAMK 698
              ||||:||:||||||||||||||||: ::| |:: |:| |||||||::||  || ::|
Db  1001 ---VLGREAARCHAYAKALHYKELEFLQDHNSGAVEALIVINNQLQQSDAAIGILRKVKT 1057

Qy   699 HFGELEIQATWYEKLHEWEDALVAY-DKKMDTNKDDP--ELMLGRMRCLEALGEWGQLH 754
            :||| :: :| ||| ||::|| :| : |: : ||:|  :::: ||||  |||||  | :
Db  1058 YREGIQLRESWFEKLERWDEALNFYCQRREREIPEDQPVPVDIVWGKMRCYHALGEWDSLA 1117

Qy   755 QQCCEKWTLVNDETQAKMARMAAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVLALHQ 814
           : || ||| |: :::| :| |:|::|||||:|||| ::  : ::| | ||| |:::||:
Db  1118 TLAGKTWANSAPEIQRRIAPLATTAAWGLGKWDSMDSYLQSMKRFSPDRAFFGAIIALHR 1177

Qy   815 DLFSLAQQCIDKARDLLDAEITAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVP-ERRE 873
           :  :  :: ||:|| || ||::|:::||| ||: |::| :|:||| | | |::   |::
Db  1178 NQFREAMACIIQAARGLDTELSALVSESYNRAYCVVVRVQMLAEELELIVYKQCDGEKQA 1237

Qy   874 IIRQIWWERLQGCCRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRIALAHKTL 933
           :: ||| |||:|||::||:|| :|: ||| ||:|::|   || :||| ||| ::||:|:|
Db  1238 TMRRTWETRLKGCCQRNVEVWQRMLRLRSLVMTPQENMEMWTKFANLCRKSGRMGLAEKSL 1297

Qy   934 VLLLGVDPSRQLDHPLPTVHPQ--------VTYAYMKQMWKSARKIDAFQEMQHFV 981
            :  |:|||  | :|:|| | :         | ::::| |: |:| |:::: :::|
Db  1298 KQLIGTD--SPLDSVIPYWHDRHPGSVGNRIASPILYAVLKFQWEIGQS-PGFRNSEHRV 1354

Qy   982 --------QTMQQAQEALAT-------------------EDQCHKQE 1002
                    :|:  | : | :                   |:: ::
Db  1355 AEKTLYCLQRFTQETAHRYETSRMHMAAHAQNGMEVQNQPGPAEFNEEMMHPQAQKHWLE 1414

Qy  1003 LHKLMARCFLKLGEW--QLNLQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVNFEA 1060
           : |: |:| ||||| | |:|:|| :: : :| ||:: :|| |:||| ||| ||: |  |
Db  1415 QTVLLAKCYLRQGEWMVSLNKDDWQQRYRAEVLDCYYKATHYNNKWYKAWHAWALLANFE- 1473

Qy  1061 VLHYKHQNQARDEKKKLRHASGANITNATTAATTAATATTTASTEGSNSESEABSTENSP 1120
            | |: |: :: |                                  |::
Db  1474 VVQYITGNREVDVR----------------------------------NNGE-------- 1491

Qy  1121 TPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRLVLTLWFDYGHWPDVN 1180
                 : :: |:|::|| | || ::| ||:::|| :|::| :||||:| |: ||::| :|
Db  1492 --------QNYIIQYVVPAVRGFFEEIALSSGGSLQDTLRLLALLMLTHGGHIDVH 1538

Qy  1181 EALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLIVAS 1240
           :|:||| :|| :||||:|||||||||: | ::::|::|::| |:||:|||||:||| |: 
Db  1539 NAVTEGFTRVSVDTWLEVIPQLIARINQPHKRVQQSVHGLLADVGRAHPQALVYPLTVAR 1598

Qy  1241 KS-TTTARHVAANKIIKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYF 1299
           ||   | : :::|:|::|| ||||||:||| ||| ||||||:||||:|||::|||||||||
Db  1599 KSWHNTRVRSANQILESMRQHSLRLVEQAEIVSGEHIRVAVLWHELWHEALEEASRLYF 1658

Qy  1300 GERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQ 1359
           | :|: ||| : |||||::| ||||::: |  |:||||| ::|:||::| :: ::|:| |
Db  1659 GDHNLEGMFATTEPLHELLEAGPQTLREISPAQTFGRDLSEARDWCRQYQETQDSNDLNQ 1718

Qy  1360 AWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPS 1419
           |||||| |||||::||||:|||||:| || |:| :|||||||||   | :: || |   
Db  1719 AWDLYYTVFRRITRQLPQMTSLELAYCSPNLLNAKDLELAVPGTYRSGQEVVRIMSFDAT 1778

Qy  1420 LQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKN 1479
            :| ||||||||| ::|| ||: |:|||||||:|||||||||||| |||||: : ::|:
Db  1779 FSVISSKQRPRKLDIVGSDGKTYTFLLKGHEDIRQDERVMQLFGLCNTLLANDSECFKRH 1838

Qy  1480 LSIQRYAVIPLSTNSGLIGWVPHCDTLHLALIRDYREKKILLNIEHRIMLRMAPDYDHLT 1539
           |:|||:|:||||::|||:|||| ||| ||::||:||::||||||||||||:|||||||:||
Db  1839 LNIQRYPAIPLSQNSGLLGWVPNSDTVHQLIREYRDSRKILLNIEHRIMLQMAPDYDNLT 1898

Qy  1540 LMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEWFDRRTNYTRSLAVMSWVGYILGLGDR 1599
           |||||||| :|::||:| ||: :|||||:||| ||||||||||| |||||||||||||||
Db  1899 LMQKVEVFGYALDNTTGQDLYRVLWLKSKSSEAWLDRRTNYTRSLGVMSWVGYILGLGDR 1958

Qy  1600 HPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPPRLITRMLTNAMEVTGLLDGNVRIT 1659
           |||||||||::||:|||||||||||||:||:|||: | ||:||:||:||||: :: ||:|:|
Db  1959 HPSNLMLDRITGKIIHIDFGDCFEVAMKREYPEKVPERVPPRLITRMLTYAMEVSNIEGSRTT 2018

Qy  1660 CHTVMEVLREHKDSVMAVLEAF------------------VYDPLLNWRLMDTNT--- 1696
           | :|| |||: |:||||||||                   ::|||||||||:: :
Db  2019 CEHVMRVLRENKESVMAVLEAVSILCVPFLTGSLIDTRNSSFHDPLLNWRLINPTSPPG 2078

Qy  1697 ----------KGNKRSRTRTDS-------------YSAGQSVELLDGVELEGPAH--KKTGT 1733
                     ||:| || |::             |||| ||  | |::|||
Db  2079 PNFNAERRVAMPGGPGPRARRPSILDAPVAPTEFLAAQAAPGADGHGMGMSANPGRSRAR 2138
```

```
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2380

Query Match        46.7%;   Score 4398;   DB 15;   Length 1973;
Best Local Similarity  48.1%;   Pred. No. 0;
Matches  888;   Conservative 324;   Mismatches 476;   Indels 158;   Gaps 20;

Qy      1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKALILKLKDPDDPNPGVINNVLA  60
           |:|:|| ||                |: :|| ||||:|::|:|
Db    249  LDYSTIIRTKEENAKLLCLLIAAAPRLIRSHVEPILQILLPKAKDSSSIVAASIVN----  304

Qy     61  TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYYVEPYRK  120
           :|::|:|:|   :|:  :|:  :|:|| :| ||::|:||||:|:|:|
Db    305  SLGEICQISGEVIVPFIKDLMPLIIEALQDQSPIRRAAALKAIGNLSSSTGYVIDPYIE  364

Qy    121  YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  180
           |  |:||:|||||    |:|  | :|||:|| |:|:  :|||:|:::
Db    365  FPSLLDLILIGITKTEQDITIRRETIKLIGTLGALDPNRHRV-  405

Qy    181  KSSQDSSDYSTSEMLVNMGNLP----------LDEFYPAVSMVALMRIFRDQSLSHH  227
                              :|| |            |:| ||:|||:||:|| ::|:|
Db    406  -----LEKGTEKVVPEQKNIPPDISLLMSGIGPSDEYPTVVITALMSILKDPSLTIH  459

Qy    228  HIMVVQAITFIFKSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHI  287
           ||   |:|||:| |:|:|  : |: |: :| ||::||| :| |::| |:|:
Db    460  HTAVIQAVMYIFKIMGLRCAPFLSQIIPEFIAVMRTCPTNILEFYFQQLSIIVLIVRQHI  519

Qy    288  RPYMDEIVTLMREFWVNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVPMHDNS  347
           | |::|  ||:| || ::| :|::| |:| |:|
Db    520  RSPLPDLFKLIKDFWNPHSNLQFTIISLIFSLARAMQGEFKPYLPSLLVMMLQIFDSDVS  579

Qy    348  PGRIVSIKLLAAIQLFGANLDDYLHLLLPIVKLFDAPEAPLPSRKAALETVDRLTESLD  407
           ||:|:|| |||  ||   :|::| |::|:| |:| |::: :
Db    580  VDSVSTKKVLHAFIVFGDTLADYFHMLLDPILRLYERNDVSIGIKESIMITIGRLSMVIN  639

Qy    408  FTDYASRIIHPIVRTLD-QSPELRSTAMDTLSSLVPQLGKKYQIFIPMVNKVLVRHRINH  466
           ||:||||||| || |:|  | |:  : ::|:| |:| |:|:: :|:|:||:|
Db    640  LSEYASRIIHPVVMRMLSCNNASLIRVSMDTVCALIYQLNVDFAITIPMIDKCLKMNGVTH  699

Qy    467  QRYDVLICRIVKGYTLADEERDPL----IYQHRMLRSGQGDALASGPVETGPMKKLHVSTI  523
            : |:||| ||| |||:||||||     |::| |:| :||:|| :||| ||:|| ||:|:|
Db    700  -----QEQPLPIKLNPYEKYDKPKLDVVASAADITS--KKLPVNQE  749

Qy    524  NLQKAWGAAARVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFV  583
           ||:|||| | :|:||||:||||||||:|:|||:||||||| |:| | |:| |:|||||||
Db    750  ILRNAWEASQRSTKDDWQEWIRRIGVALLRESPSHALRACAALAAAYQPLARDLFNASFV  809

Qy    584  SCWSELNEDQQDELIRSIEIALTSQDIA--EVTQTLLNLAEFMEHSDKGPLPLRDDNGIVL  642
           ||||||:|:||:||:||:| |::| :     ||:| | :|||:|| | |:|   |:|:|:
Db    810  SCWSELYDHFQEELVKSIEIALTSPHISPEIIQLLNLAEFMEHDDK-PLPI----DIRT  864

Qy    643  LGERAAKCRAYAKALHYKELEFQK-----GPTPAILESLLSINNKLQQPEAAAGVLEYAM  697
           || |  | :||:||||:||:|::     :|:||:||::|:::|| :| :|||::::|
Db    865  LGAYAAKCHAFAKALHYKELEFIEEELVTKPSVDTIEALLISINNQLQQPDAAIGILKHAQ  924

Qy    698  KHFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMCLELALGEWGQLHQQC  757
            |||| || |::|||||:|||:|::||:: :|::|  ||: ||::|:|:|| ||| |:|:
Db    925  QH-DKMNLKETWYEKLQRWEDALSAYEKREAAGAGNFEITMGKLRCLHALGEWDRLSQLA  983

Qy    758  CEKWTLVNDETQAKMARWAAAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLF  817
              :| | |:|  | ||: |::   |:|| ||::||:|| | :  |:|| ||:||:|::|
Db    984  QENWIHAGHDARRYIAPLSVAAAWGLGQWEQMDEYISVMKSESPDKAPFNAIVALHRSQF  1043

Qy    818  SLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLV---PERREI  874
            |: | |:|| :||  | :|  :|  | ||::| |:| |:| |::|:|:|     |:|||
Db   1044  EEAASYITRARDLLDTELITALVGESYNRAIYSVAVRVQMLSELEHIITYKKAEDKPEVREM  1103

Qy    875  IRQIWWERLQGCQRIVEDWQKILMVRSLVVSPHEDMETWLKYASLCGKSGRIALAHKTLV  934
           |:: |:|||:|| | | :| |::|  |: |:|  | | |||:|: ||:|| ||::: :|
Db   1104  IKKTWVRRLKGCQRNVDVWQRMLRIRSLVISPRDNMENWIKFANLCRKSGRISLAEKSLN  1163

Qy    935  LLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQ--------  986
           ||| |  |  :  :|                |  :|
```

```
Db   1164  LLEDDENLDNSLVLKKTHPSIVYANLKFPLWAVDDKRKALNSMQEFTSQLISDINVPDAL  1223
           |:|:|:|| |:|||:|||:|| | |||| ||  |||:|||||||||: ::|::|| ||:
Qy    987  -QAQHAIATEDQQHKQELHKIMARCFLKLGEWQLNLQ-GINESTIPKVLQYYSAATEHD  1043

Db   1224  FVQSTSVNTKSQEEIQYYFHLLARCYHKQGQWQQEIENNWSEGSFDGVLQSYMYATQPD  1283
           |:| || :|||:|| || |:|::|  || |:|: |:| ::| ||:| :|:| |:|: |
Qy   1044  RSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANINATTAATAITTAS  1103

Db   1284  SKWYKAWHSWALANFEAVKFLE--------  1305
            |||||| | |  ||||
Qy   1284 (blank)

Db   1104  TEGSNSESRAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDT  1163
           : :  :     ::   |:|   :|:|                |::|:|:| :  ||:||
Qy   1306  ------------QSEQIPSAAYEQ-------------YIIPAVKGFFKSIALSKG-NLQDT  1341

Db   1164  LRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTD  1223
           |::||:| |  |  | :||| :|| |||:| ||:|:|||:|| ||::|| |:|:|:|||:
Qy   1342  LRLLNLWFKFGNNSNVINTLNVGISTVANIDIWLDVIPQLIARIHASSLSVRKSVHQLLSD  1401

Db   1224  IGRYHPQALIYPLTVASKSTTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVALIW  1283
           :||:||||:| |||| |:||:  | || |:|:|::  |::||:||||: :||| |:|:|
Qy   1402  VGRAHPQALVYPLTVAAKSQSSARQNAALAIMDSLKTHSPRLVEQARLVSHELIRAAILW  1461

Db   1284  HEMWHHGLEEASRLYFGERNVKGMPEVLEPLHAMMERGPCTLKETSFNQAYGRDLMEACE  1343
           |:| || |||||||||| :| ||:|| || ||  ||||  |:| |:| | || || |
Qy   1462  HEQWHEGLEEASRLYFGGDHNIEGMEAVLRPLHEMLERGPETLREISFQQAFGRDLVEARD  1521

Db   1344  WCRKYMKSGNVKDLTQAWDLYYHVRRISKQLPQLTSLELQYYSPKLLMCRDLELAVPGT  1403
           || | | ||:|:| ||||| |:|:| ::|  ||:||::||::|||:| || |:|||||
Qy   1522  CCIRFEQTGDISDLNQAWDLYQVFKKIRKQLPQLTTLDLQYYSPKLLHVHDLELAVPGT  1581

Db   1404  YPDNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFG  1463
           | ::||:| | ::  |  |:||:||||||:| |||| |||:||||||: || |||:|||:
Qy   1582  YVGSKPVIRIVKFYPTFNVITSKQRPRRLSIKGSDKDYQVVLKGHEDIRQDERVMQLFG  1641

Db   1464  LVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYRKKKILLNI  1523
           | |:|||:|||::|:||||||  | |:  | ||||||  |||| |||||||:::|||:|
Qy   1642  LCNNLLLADPETFKRLLSIQRYPVIPLSPDSGLLGWVLDSDTLHVLIRDYRSRKLLLNI  1701

Db   1524  EHRIRLRMAPDYDHLTLMQKVEVFEHAVNNTAGDLLAKLLWLKSPSSEVWFDRRTNYTRS  1583
           ||||:|::|||| |:||::||||||:||::| || |:|:||||||| :|: |: || : |
Qy   1702  EHRLIIQMAPDYRLTLLQKVEVFRYALLSTTGQDLYRVLMLKSRSSAAWLNRRTNYSRT  1761

Db   1584  LAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFRVAMVTREKPPEKIPFRLTRML  1643
           |||||::||:|||||||| |||||| |||:||||||||| ||  |:|| ||||||||||
Qy   1762  LAVMSMVGYIIGLGDRHPSNLMLDRYTGNIIHIDFGDCFEVAMHREKPPEKIPFRLTRML  1821

Db   1644  TNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLINWRLMDTNTKGNKRSR  1703
           | ||| |:|    ::|  ||:|::| ||:||:||:|| |||||||:: |::::||: :|
Qy   1822  VNAMEVSGIEGTFRITCEHVMRVLRTNKESVMAVLEAFVVDPLINWRLAPAYSPSIDEKQ  1881

Db   1704  TRTDSYSAGQSVEILDGVELGRPAHKKTGTTVPESIHSFIGDGLVKPERALNKKAIQIINR  1763
            : ::::| || |:||:|| |:   |:  ::| |:|:||||| |:| ||  :|| |::|:
Qy   1882  SNEPNTLLGETIDGLHRKRLNE------EGITLEER--------QKPEILNQRAITVLNR  1927

Db   1764  VRDKLTGRDFSHDDTLDVPTQVELLLIKQATSHENLCQCYIGWCPFW  1809
           | |||:||| | |:||:||:||  |:|:||| :|||| |||||| ||
Qy   1928  VSNKLTGRDFKPQQQLDVPSQVEKLILQATSIRNLCLCVIGWCSFW  1973
```

RESULT 5
US-10-369-493-3923
; Sequence 3923, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

```
Qy  1795  HENLCQCYIGW 1805
          |||||||||||
Db  1812  HENLCQCYIGW 1822
```

RESULT 3
US-09-950-634-4
; Sequence 4, Application US/09950634
; Publication No. US20030032775A1
; GENERAL INFORMATION:
; APPLICANT: Molnar-Kimber, Katherine L.
;   Failli, Amedeo F.
;   Caggiano, Thomas J.
;   Nakanishi, Koji
;   Chen, Yanqiu
; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;   Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/950,634
; FILING DATE: 13-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,112
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/384,524
; FILING DATE: 13-FEB-1995
; APPLICATION NUMBER: US 08/312,023
; FILING DATE: 26-SEP-1995
; APPLICATION NUMBER: US 08/207,975
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Siekman, Michael T.
; REGISTRATION NUMBER: 36,276
; REFERENCE/DOCKET NUMBER: 01142.0058-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-950-634-4

Query Match   49.8%; Score 4690; DB 10; Length 1140;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 894; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
Qy   915  KYASLCGKSGRIALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTTYAYMKQMWKSARKIDAF  974
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   246  KYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTTYAYMKQMWKSARKIDAF  305

Qy   975  QHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ  1034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306  QHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ   365

Qy  1035  YYSAATEHDRSWYKAWHAWAVMNFEAVLEYKHQNQARDEKKKLRHASGANITNATTAATT  1094
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366  YYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANTNATTAATT    425

Qy  1095  AATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL  1154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426  AATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL   485

Qy  1155  SRGNNLQDTLRVILTLWFDYGHWPDVNEALVBGVKAIQIDTWLQVIPQLIARIDTPRPLVG 1214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486  SRGNNLQDTLRVILTLWFDYGHWPDVNEALVBGVKAIQIDTWLQVIPQLIARIDTPRPLVG  545

Qy  1215  RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSE   1274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546  RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSE    605

Qy  1275  ELIRVAILWHEMWHEGLEPASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY  1334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606  ELIRVAILWHEMWHEGLEPASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY   665

Qy  1335  GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCR  1394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666  GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCR   725

Qy  1395  DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKITLMGSNGHEFVFLLKGHEDLRQ  1454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726  DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKITLMGSNGHEFVFLLKGHEDLRQ   785

Qy  1455  DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR  1514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786  DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR   845

Qy  1515  EKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWF  1574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846  EKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWF   905

Qy  1575  DRRTNYTRSLAVWSMVGYILGLGDRHPSNLMLDRLSGKLIHIDFGDCFEVAMTREKFPEK  1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906  DRRTNYTRSLAVWSMVGYILGLGDRHPSNLMLDRLSGKLIHIDFGDCFEVAMTREKFPEK   965

Qy  1635  IPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT  1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966  IPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT  1025

Qy  1695  NTKGNKSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN  1754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1026  NTKGNKSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN  1085

Qy  1755  KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW      1809
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1086  KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWYPFW      1140
```

RESULT 4
US-10-369-493-2380
; Sequence 2380, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2380
; LENGTH: 1973
; TYPE: PRT

RESULT 2
US-10-334-143-25
; Sequence 25, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 1822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-25

Query Match   86.7%; Score 8164.5; DB 15; Length 1822;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 1601; Conservative 4; Mismatches 4; Indels 205; Gaps 5;

```
Qy     1 LEHSGIGRIKEQSARMLGHIVSNAPRLIRPYMEPILKALILKDPPDPNPGVINNVLA   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211 LEHSGIGRIKEQSARMLGHIVSNAPRLIRPYMEPILKALILKDPPDPNPGVINNVLA  270

Qy    61 TIGELAQVSGLEMRKVVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   271 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK  330

Qy   121 YPTLLEVLLNFLKTEQNQCTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   331 YPTLLEVLLNFLKTEQNQCTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  390

Qy   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   391 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK  450

Qy   241 SLGLKCVQFLPQVMFTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   451 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE  510

Qy   301 FWTMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVEMHDNSPGRIVSIKLLAAI  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   511 FWTMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVEMHDNSPGRIVSIKLLAAI  570

Qy   361 QLFGANLDDYLHHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   571 QLFGANLDDYLHHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV  630

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   631 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY  690

Qy   481 TLADEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQK-----AWGAARR  534
           ||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db   691 TLADEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKCASPLFQAWGAARR  750

Qy   535 VSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQ  594
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   751 VSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQ  810

Qy   595 DELIRSIELALTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYA  654
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   811 DELIRSIELALTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYA  870

Qy   655 KALHYKELFFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELBIQATWYEKLH  714
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   871 KALHYKELFFQKGPTPAILESLIR------------------------------------  894

Qy   715 EWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKWAR  774
                                        |||||||||||||||||||||||
Db   895 ----------------------------GQLHQQCCEKWTLVNDETQAKWAR  918

Qy   775 MAAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAE  834
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   919 MAAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAE  978

Qy   835 LTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQ  894
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   979 LTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQ  1038

Qy   895 KILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHP  954
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1039 KILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHP  1098

Qy   955 QVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKL  1014
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 QVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATBDQQHKQELHKLMARCFLKL  1158

Qy  1015 GEWQLNLQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEK  1074
           |||||||||||||||||||||||||||||||||||
Db  1159 GEWQLNLQGINESTIPKVLQYYSAATEHDRSWYK--------------------------  1192

Qy  1075 KKLRHASGANITNATTAATANTATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLS  1134
                                                                    |||
Db  1193 ---------------------------------------------------------DLS  1195

Qy  1135 KTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDT  1194
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1196 KTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDT  1255

Qy  1195 WLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKI  1254
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1256 WLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKI  1315

Qy  1255 LKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWEBGLEEASRLYFGERNVKGMFEVLEPL  1314
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1316 LKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWEBGLEEASRLYFGERNVKGMFEVLEPL  1375

Qy  1315 HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQ  1374
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1376 HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQ  1435

Qy  1375 LPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTL  1434
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1436 LPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTL  1495

Qy  1435 MGSNGHEVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNS  1494
           |||||||||||||||||||||||||||||||||||||||||||
Db  1496 MGSNGHEVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLS-------------  1542

Qy  1495 GLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNT  1554
                                           |||||||||||||||||||||||||
Db  1543 ------------------------------C------VLMQMAPDYDHLTLMQKVEVFEHAVNNT  1571

Qy  1555 AGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSWVGYILGLGDRHPSNLMDRLSGKIL  1614
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1572 AGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSWVGYILGLGDRHPSNLMDRLSGKIL  1631

Qy  1615 HIDFGDCFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSV  1674
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1632 HIDFGDCFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSV  1691

Qy  1675 MAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTT  1734
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1692 MAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTT  1751

Qy  1735 VPESIHSFIGDGLVKPEALNKKAIQINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATS  1794
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1752 VPESIHSFIGDGLVKPEALNKKAIQINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATS  1811
```

```
ATTORNEY/AGENT INFORMATION:
   NAME: Siekman, Michael T.
   REGISTRATION NUMBER: 36,276
   REFERENCE/DOCKET NUMBER: 01142.0058-00000
TELECOMMUNICATION INFORMATION:
   TELEPHONE: 202-408-4000
   TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 3:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 2549 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
   MOLECULE TYPE: peptide
   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-950-634-3

Query Match       99.9%;    Score 9402;  DB 10;  Length 2549;
Best Local Similarity  99.9%;   Pred. No. 0;
Matches 1808; Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA 800

Qy      61 TIGELAQVSGLEMRKWVDELFPIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     801 TIGELAQVSGLEMRKWVDELFPIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 860

Qy     121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     861 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy     181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy     241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy     301 FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMDNSPGRIVSIKLLAAI 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1041 FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMDNSPGRIVSIKLLAAI 1100

Qy     361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLSRKAALETVDRLTESLDFTDYASRIIHPIV 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1101 QLFGANLDDYLHLLLPPIVKLFDAPEAPLSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy     421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy     481 TLADEERDPLIYQHRMLRSGGQDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1221 TLADEERDPLIYQHRMLRSGGQDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy     541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy     601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy     661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1401 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy     721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy     781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTPAMAG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTPAMAG 1580

Qy     841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 1640

Qy     901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy     961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATESDQQHKQELHKLMARCFLKLGEWQLN 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy    1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAWMNFEAVLHYKHQNQARDEKKKLRHA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAWMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy    1081 SGANITNATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1821 SGANITNATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880

Qy    1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy    1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 2000

Qy    1261 HSNTLVQQAMMVSEELIRVAILWEEWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2001 HSNTLVQQAMMVSEELIRVAILWEEWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy    1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy    1381 LELQVYSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2121 LELQVYSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy    1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy    1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy    1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy    1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy    1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy    1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy    1801 CYIGWCPFW 1809
           ||||| ||||
Db    2541 CYIGWYPFW 2549
```

```
                      GenCore version 5.1.6

             Copyright (c) 1993 - 2004 Compugen Ltd.


OM protein - protein search, using sw model

Run on:     March  2, 2004, 20:03:28 ; Search time 51 Seconds
                                       (without alignments)
                                       7489.730 Million cell updates/sec

Title:           US-09-517-491-12
Perfect score:   9413
Sequence:        1 LEHSGIGRIKQSARMLGHL.........KQATSHENLCQCYIGWCPFW 1809

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters:       809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Published_Applications_AA:*
          1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
          2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
          3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
          4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
          5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
          6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
          7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
          8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
          9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
         10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
         11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
         12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
         13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
         14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
         15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
         16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
         17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
         18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result           %      Query
No.     Score  Match  Length  DB  Length  ID                     Description
-----  ------  -----  ------  --  ------  ---------------------  -----------
  1     9402   99.9    2549   10  US-09-950-634-3                Sequence 3, Appli
  2    8164.5  86.7    1822   15  US-10-334-143-25               Sequence 25, Ap
  3     4690   49.8    1140   10  US-09-950-634-4                Sequence 4, Appli
  4     4398   46.7    1973   15  US-10-369-493-2380             Sequence 2380, Ap
  5     4155   44.1    2222   15  US-10-369-493-3923             Sequence 3923, Ap
  6     4147   44.1    2335   15  US-10-369-493-2561             Sequence 2561, Ap
  7     4105   43.6    2473   15  US-10-369-493-22217            Sequence 22217, A
  8     3962   42.1    2470   15  US-10-369-493-22191            Sequence 22191, A
  9     3913   41.6    2470   15  US-10-310-154-579              Sequence 579, App
 10     3417   36.3    2697   15  US-10-369-493-5036             Sequence 5036, Ap
 11     842.5   9.0    2539   15  US-10-369-493-3779             Sequence 3779, Ap
 12     753.5   8.0    2644   14  US-10-300-453A-39              Sequence 39, Appl
 13     734.5   7.8    2654   14  US-10-227-610-2                Sequence 2, Appli
 14     729     7.7    3521   15  US-10-165-216-2                Sequence 2, Appli
 15     726     7.7    3529   15  US-10-165-216-8                Sequence 8, Appli
 16     726     7.7    3657   15  US-10-165-216-10               Sequence 10, Appl
 17     725.5   7.7    2386   15  US-10-369-493-2379             Sequence 2379, Ap
 18     693     7.4    2930   10  US-09-957-837A-2               Sequence 2, Appli
 19     693     7.4    3031   15  US-10-165-216-6                Sequence 6, Appli
 20     634     6.7    3056   14  US-10-175-225-2                Sequence 2, Appli
 21     631     6.7    3056    8  US-08-984-090-2                Sequence 2, Appli
 22     631     6.7    3056   10  US-09-825-476-3                Sequence 3, Appli
 23     631     6.7    3056   14  US-10-307-077-1                Sequence 1, Appl
 24     631     6.7    3056   14  US-10-059-585-41               Sequence 41, Appl
 25     631     6.7    3056   14  US-10-361-562-2                Sequence 2, Appli
 26     631     6.7    3056   14  US-10-116-275-338              Sequence 338, App
 27     630.5   6.7    3057   14  US-10-175-225-3                Sequence 3, Appli
 28     630     6.7    230    15  US-10-341-967-111              Sequence 11, App
 29     628     6.7    2368   15  US-10-369-493-1468             Sequence 1468, Ap
 30     552.5   5.9    4128   14  US-10-205-194-1                Sequence 1, Appli
 31     544     5.8    129    15  US-10-341-967-79               Sequence 79, Appl
 32     505     5.4    128    14  US-10-270-223-12               Sequence 12, Appl
 33     505     5.4    240    14  US-10-270-223-10               Sequence 10, Appl
 34     505     5.4    247    14  US-10-270-223-8                Sequence 8, Appli
 35     438     4.7    104    11  US-09-864-408A-7020            Sequence 7020, Ap
 36     413     4.4    140     9  US-09-902-180-6                Sequence 6, Appli
 37     261     2.8    887     9  US-09-771-161A-192             Sequence 192, App
 38     261     2.8    887     9  US-09-771-161A-193             Sequence 193, App
 39     227.5   2.4    873    15  US-10-369-493-5060             Sequence 5060, Ap
 40     226     2.4    875    15  US-10-369-493-1795             Sequence 1795, Ap
 41     201.5   2.1    1044   15  US-10-027-591-2                Sequence 2, Appli
 42     201.5   2.1    1044   14  US-10-162-160-1                Sequence 1, Appli
 43     201.5   2.1    1044   14  US-10-337-192-2                Sequence 2, Appli
 44     196.5   2.1    1068   14  US-10-101-235A-8               Sequence 8, Appli
 45     193.5   2.1    928    15  US-10-369-493-3211             Sequence 3211, Ap
```

ALIGNMENTS

```
RESULT 1
US-09-950-634-3
; Sequence 3, Application US/09950634
; Publication No. US20030032775A1
; GENERAL INFORMATION:
;    APPLICANT: Molnar-Kimber, Katherine L.
;               Failli, Amedeo F.
;               Caggiano, Thomas J.
;               Nakanishi, Koji
;               Chen, Yanqiu
;  TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
;  NUMBER OF SEQUENCES: 23
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;                Dunner, L.L.P.
;     STREET: 1300 I Street, N.W.
;     CITY: Washington
;     STATE: DC
;     COUNTRY: USA
;     ZIP: 20005-3315
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/950,634
;     FILING DATE: 13-Sep-2001
;     CLASSIFICATION: <Unknown>
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/471,112
;     FILING DATE: 06-JUN-1995
;     APPLICATION NUMBER: US 08/384,524
;     FILING DATE: 13-FEB-1995
;     APPLICATION NUMBER: US 08/312,023
;     FILING DATE: 26-SEP-1995
;     APPLICATION NUMBER: US 08/207,975
;     FILING DATE: 08-MAR-1994
;
```

Qy   1759 QIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
             ::|:|:||||||| |||||||:|||||:||:|||||||| |||||:
Db   512 LVLKRITDKLTGNDIKRLRGLLDVPTQVDKLIQQATSVENLCQHYIGWCSCW 562

Search completed: March 2, 2004, 20:04:40
Job time : 36 secs

;
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/360,144
; FILING DATE: 20-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: APV-036.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-012-515A-14

Query Match   18.4%; Score 1735; DB 3; Length 562;
Best Local Similarity 55.8%; Pred. No. 3e-157;
Matches   330; Conservative 90; Mismatches 129; Indels 42; Gaps 4;

```
Qy   1232 LIYPLTVASKSTTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGL 1291
          :|:||||: ||  ::|  ||| ::: | ||::   ::||  |||||||||  |||:|||:
Db      1 LVYPLTVAITSESTSRKKAAQSIIEKMRVHSPSLVDQAELVSRELIRVAVLWHEQWHDAL 60
Qy   1292 EEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKS 1351
          |:||| :||| |.|:||| ||||||| | :|||::|::| ||| | | || :||| :  :
Db     61 EDASRFFFGEHNTEKMFETLEPLHQMLQKGPETMREQAFANAFGRELTDAYEWVLNFRRT 120
Qy   1352 GNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPII 1411
          :::|||.||||:||:|||||||:| ||:||||||||| :|: :|||||||||||:   :|
Db    121 KDITNLNQAWDIYYNVFRRVSKQVQLLASLELQYVSPDLEHAQDLELAVPGTYQAGKPVI 180
Qy   1412 RIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLAN 1471
          || .:||:  |: |||||||| :  |:|:|  ::||||||:|||: ||||||||||||:|
Db    181 RIIKFDPTFSIISSKQRPRKLSCRGSDGKDYQYALKGHEDIRQDNLVMQLFGLVNTLLVN 240
Qy   1472 DPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRM 1531
          ||  :|::| |||:|:|||| : ||||||| |  ::|||::| || :|:||||||::|:|
Db    241 DPVCFKRHLDIQQYPAIPLSPKVGLLGWVPNSDTFHVLIKGYRESRSIMLNIEHRLLLQM 300
Qy   1532 APDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKPSSEVWFDRRTNYTRSLAVMSMVG 1591
          |||||:|||:|||||| ::::: ||:||  :|||| ||| |||| || |||||||||||
Db    301 APDYDFLTLLQKVEVFTSAMDNCKGQDLYKVLWLKSKSSEAWLDRRTYTYTRSLAVMSMVG 360
Qy   1592 YILGLGDRHPSNLMLDRLSGKILHIDFPGDCFEVAMTREKPPEKIPFRLTRMLTNAMEVTG 1651
          ||||||||||||||||| :||::|||||||||| | :||: |::|||||||||::|||::|
Db    361 YILGLGDRHPSNLMLDRITGKVIHIDFPGDCFEAAILREKYPERVPFRLTRMLNYAMEVSG 420
Qy   1652 LDGNYRITCHTMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNYLKGNKRSRTRTDSYSA 1711
          ::|::|||| :|:|::|.:.:|||:||||.|||| ||  ||:    :|          |
Db    421 IEGSFRITCEHVMRVLRDNKESLMAILEAFAYDPLINWG-FDFPTKALAES--------- 470
Qy   1712 GQSVEILDGVELGEPAHKKTGTTVPE-SIHSFIGDGLV---------------KPEALNKKAI 1809
                          ::::|:|   |:  :|||:|||:|               || ||::| |
Db    471 ---------------TGIRVPQVNTABLLRRGQIDEKEAVRLQKQNELEIRNARAA 511
Qy   1759 QINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
          .|  :.|| || . . |  ||||||| :||:|||.||||| |||||  |
Db    512 LVLKRITDKIITGNDIKRLRGLDVPTQVDKLIQQATSVENLCQHYIGWCSCW 562
```

RESULT 15
US-08-360-144A-14
; Sequence 14, Application US/08360144A
; Patent No. 6150117
; GENERAL INFORMATION:
; APPLICANT: Berlin, Vivian
; APPLICANT: Chiu, Maria Isabel
; APPLICANT: Cottarel, Guillaume

```
Qy    61 TIGELAQVSGLEMRKKVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK  120
         ::| :||   | | : |||:|: ||:::||| ||::   |||:|| ||||  |:|:| |
Db   774 TIGELSVVGGEDMKIYLKDLPLIIKFQDGSNSFKREAALKALGQLAASSGYVIDPLLD    833

Qy   121 YPTLLEVLLNFLKTEQNGGTRREAIRVLGLIGALDPYHKVNIGMIDQSRDAGAVSLSES   180
         || ||  :|| |||::| : ||| | |||:||| |||  |:||:||           :
Db   834 YPELLGILVNILKTENSQNIRRQTVTLIGIILGAIDPYRQK---------ERRVTSTTDI  883

Qy   181 KSSQDSSDYSTSEMLVNMGNLPL-DEFYPAVSMVALMRIFRDQSLSHHHTMVVQATFIF   239
         ::  :   |  :::| ||  | |  | |   |:  :: ||  |||  ||  |:||::|
Db   884 STEQNAPPIDIA--LLMQGMSPSNDEYYTTVVIHCLLKILKDPSLSSYHTAVIQAIMHIF  941

Qy   240 KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR   299
         ::||||||.||.|::||:|:|:|||   |||:||||| |:  |:: |||| ||:| ||:
Db   942 QTLGLKCVSFLDQIIPTILDVMRTCSQSLLEFYFQQLCSLIIIVRQHIRPHVDSIFQAIK   1001

Qy   300 EFWVMNTSIQSTILLIEQIVALGGEKKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAA    359
         :|   | :::| ||:: :||: |  :|: ||||  |:|:: ::: |  :::| ::|:
Db  1002 DFSSV-AKLQITLVSVIEAISKALEGEFKRLVPLTLTLFVILENDKSSDKVLSRRVLRL   1060

Qy   360 IQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESSLDFTDYASRIIHPI  419
         : |||.||: | :|: ||||::  ::  | |:: ::  |  :|| ::||| |:::|| :
Db  1061 LESFGPNLEGYSHLIITPKIVQMAEFTSGNL--QRSAIITIGKLAKDVDLFEMSSRIVHSL  1118

Qy   420 VRTLDQ--SPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIV   477
         :| | :   .|:||  | ||| |  |:|.   ||:| :: ||::   :|  |||:| ||:
Db  1119 LRVLSSTTSDELSKVIMNTLSLLLIQMGTSFAIFIPVINEVLMKKHIQHTIYDDLINRIL   1178

Qy   478 KGYTLLADBEEDPLIYQHRMLRSGCGDALASGPVETGP--MKKLHVSTINLQKAWGAARRV   535
                :   :::||:||  ::    | | ::|  | |   :||||:|||||:|||: ::
Db  1179 NNDVLPT----------KIEAANTDYKPAEQMERADAGVAKLPINQSVLKSAWNSSQQR    1228

Qy   536 SKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQD   595
         :|:|| ||::||:::|||:|.| :||:| ||:  | | |::||| || :||:||  :|:|
Db  1229 TKEDWQEWSKRLSIQLLKESPSHALRACSNLASMYYPLAKELFNTAFACVWTELYSQYQE   1288

Qy   596 ELIRSIEIALTSQ-DIAEVIQTLINLAEFMEHSDKG-PLPLRDDNGIVLLGERAAKCRAY   653
         :|| ::  ||:|  ||| ::|||:||.|:||| ||  |:|   ::|:: |||:| ||:|
Db  1289 DLIGSLCIALSSPLNPPEIHCTLLNLVERMEHDDKALPIPTQS-------LGEYAERCHAY  1342

Qy   654 AKALHYKELEFQKGPTPAILLESLIISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKL   713
         |||||||| :| |.   | ||||:||||||:|:|  ||:||||||||||||| ||||:|||
Db  1343 AKALHYKEIKFIKEPENSTIESLISINNQLNQTDAAIGILKHAQQEH-SIQLKETWFEKL   1401

Qy   714 HEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMA   773
         | ||||:||:::: .:..||.|:|| :|| ||||||||:| ||| :|  ||| :||||
Db  1402 ERWEDALHAYNERERKAGDTSVSVTLGKMRSLHALGEWEQLSQLAARWKVSKLQTKKLLA   1461

Qy   774 RMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVIALHQDLFSLAQQCIDKARDLLDA   833
         ::|| | |||| |||:| | |:|  :  |.|: ||:| || ||  :|  |:|||||||:|
Db  1462 PLAAGARWGLGEWDMLEQYISVMKPKSPDKEFFDAILYLHKNDYDNASKHILNARDLLVT   1521

Qy   834 ELTAMAGESYSRAYGAMWVSCHMLSELEBVIQYKLVP---BRRBIIRQIWWERLQGCQRIV   890
         |: |:|.| :||: ||:||::::||| | :|||||:|   :||:||::| :|||:||:::
Db  1522 EISALINESTNRAYSVIVRTQIITFEEIIKYKQLPPNSBKKLHYQNLWTKRLLGCQKNV   1581

Qy   891 EDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLL--GVDPSRQLDHP   948
         : ||::| :|||::.|: : ::|:|:| || :|||: ||||::|:||    :::
Db  1582 DLWQRYLRVRSLVIKPKPKQDLQIWIKFANLCRKSGRMRLANKALNMLLEGGNDPS------   1635

Qy   949 LPT---VHPQVTYAYMKNMWKSARKIDAFQHMQHFV----------QTMQQQAQHAIAT   994
         || :    |:|  | :| ||::| || ::: :: ::          ::: ||:: ::
Db  1636 LPNTVKAPPPVVVAQLKYIWATGAYKEALNHLIGFTSRLAHDLGLDPNNMIAQSVKLSSA   1695

Qy   995 EDQQHKQELHLHARCFLKLGEWQLNLQGINESTIP-KV-QYYSAATEHDRSWYKAWIAW   1053
         .   :::| ||:|||||:||| :: ::::||: :|  ||  ||::| ::::||||| :|
Db  1696 STAPYVEBYTKLLARCFLKQGEWRIATQPNWRVINPDAIIGSYLLATHIFDKNWYKAWHNW   1755

Qy  1054 AVMNFEAVLHLYKHQNQARDEKKKLRHASGANTTNATTAAATTAATTATTTASTEGSNSESEA   1113
         |: ||| : ::  :   |:::::|
Db  1756 ALANFEVI------SMVQEETKLN--GGKNDDDDDTAVN-----NDNVRIDGSILGSSS    1801

Qy  1114 ESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVGGFPRSISLSRGNNLQDTLRVLTLWFDY   1173
         .| ::   | | :::: ||:|  : || :||::||||||:||:||||||||||:|||||
Db  1802 LTINGNRYPLELIQR---------HVVPAIKGFFFSISLLETSCLQDTLRLLTLLFNF   1850

Qy  1174 GHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALI   1233
         |:  :|::|: |: |::||| :|:::||||:||  | :|| ||: || :|| ::|||||
Db  1851 GGIKEVSQAMYEGFNLMKIENWLEVLPQLISRIHQPDTVSNSLLSLLSDLGKAHPQALV   1910

Qy  1234 YPLTVASKSTTTARHNAANKILKNMCEHSNTLVQQAMVSEBLIRVAIIWHEMWHEGLEE   1293
         ||||||:|   ::|: ||  : :: | ||  |:||| ||:| |||||::|| :| |||:
Db  1911 YPLTVAIKSEBSVSRQKAALSIIEKIRIHSPVLNQAEIVSHELIRVAVLWHELWYEGLED   1970

Qy  1294 ASRLYFGERNVKGMFEVLEPLHAMMERGPCTLKETSFNQAYGRDLMEAQEWCRKYMKSGN   1353
         |:: :|.|.|::.|| ||||||::: .. :||:|.||::::||||::|:||:|:|:||:
Db  1971 ARRQFFVEHNIEKMFSTLEPLHKHLGNEPQTLSEVVSFQKSFGRDLNDAYEWLNNYKKSKD   2030

Qy  1354 VKDLTQAWDLYYHVFRRISKQLPQLTSLBLQYVSPKLLMCRDLBLAVPGTYDPNQPIIRI   1413
         ::: :|||| ||:||||:::|:|||  | ||:||| || :::| |||||||  :|.||||
Db  2031 INNLNQAWDIYYNVFRKITRQIPQLQTLDLQHVSPQLLATHDLELAVPGTYPGKPTIRI   2090

Qy  1414 QSIAPSLQVITSKQRPRKUTLMGSNGHEFVPLLKGHEDLRQDERVMQLFGLVNTLLANDP   1473
         ::: | ::||:||||||| :|:| :|:: :||||||||:|||: ||||||||||||:||:
Db  2091 AKFEPLFSVISSKQRPRKFSIKGSDGKDYKYVLKGHEDIRQDSLVMQLFGLVNTLLKNDS   2150

Qy  1474 TSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYEEKKKILLNIEHRIMLRMAP   1533
           |: :|.||:|:.|||| :||||||||: ||:| || |::|:| | ||| :::: :||
Db  2151 BCFKRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHRDAKKLPLNIEQWVMLQMAP   2210

Qy  1534 DYDHLTLMQKVEVPEHAVNNTAGDDLAKLLWLKSPSSEVWFDRKTNYTRSLAVMSMVGYI   1593
         |||:|||:|| || : ||:||.|:||.||:|||| |||.| :|:|:||||||||| |||
Db  2211 DYENLTLLQKIEVFTYALDNTKGQDLYKLIWLKSRSSETWLERRTYTRSLAVMSMTGYI   2270

Qy  1594 LGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPPEKIPFRLTRMLTNAMEVTGLD   1653
         |||||||||||||||::||::|||||||||.|: |||.||:||||||||| ||||:|::
Db  2271 LGLGDRHPSNLMLDRITGKVIHIDFGDCFEAAILREKYPEKVPFRLTRMLTYAMEVSGIE   2330

Qy  1654 GNYRITCHTVMEVLREHKDSVMAVLRAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQ   1713
         |.::|||.:|| ||| :|:|:|| |.||. ||| :| :: .:||.     | ::  :|:
Db  2331 GSFRITCENVMRVLRDNKESLMAILEAFALDPLIHWGFDLPPQKLTEQT----------GI   2381

Qy  1714 SVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDF   1773
            ::  :   .|||: .: :|:|::. | .  |::::   :|:|:|:: |::|||||:|:
Db  2382 PLPLINPSEL---LRKGAITVEEAANM---EARQQNETRNARAMLVLRRITDKLTGNDI   2434

Qy  1774 SHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW   1809
             ||||:||:||:|:||||| |||||.||||||||
Db  2435 KRFNELDVPEQVDKLIQQATSIERLCQHYIGWCPFW   2470
```

```
Db  1229  TKEDWQEWSKRLSIQLLKESPSHALRACSNLASMYYPLAKELFNTAFACVWTELYSQYQE  1288
          :.||.||    .|.||  ||:::::|       |||||  |           :|: :
Qy   596  ELIRSIELALTSQ-DIAEVTQTLLNLABFMEHSDKG-PLPLRDDNGIVLLGERAAKCRAY   653
          .:.:|| :|| :  ::     |:.|| :::||
Db  1289  DLIGSLCIALSSPLNPPBIHQTLNLNIVEFMEHDDKALPIPTQS------LGSYABRCHAY 1342

Qy   654  AKALHYKELEFQKGPTPAILESLLSINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKL   713
          |||||||| |||:|  |:|||:||||||:|:|   || :||::|||||||:|.|:||:|||
Db  1343  AKALHYKEIKFIKEPENSTIESLLSINNQLNQTDAIGILKHAQQHI-SLQLKETWFEKL  1401

Qy   714  HEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMA   773
          ..||||: |: |:  .|  | ||:||:||:| ||||| || |  .:||.:|: : .: :|
Db  1402  ERWEDALHAYNEREBKAGDTSVSVTLGKMRSLHALGEWEQLSQLAARKWKVSKLQTKKLIA 1461

Qy   774  RMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDA   833
          .:|| .||||||.||::||..::.| .:|.||: |::|||:|:|.:::: |||||||:|.|
Db  1462  PLAAGARWGLGEWDMLEQYISVMKPKSPDKEFFDAIILYLHKNDYDNASKHIINARDLIVT 1521

Qy   834  ELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVP--ERREIIRQIWWERLQGCQRIV   890
          |:: :.||:|:||| ::|  ::||:.||:|:||:|| .|:..|||.|  :||.||:::|
Db  1522  EISALINESYNRAYSVIVRTQIITEFEEIIKYKQLPPNSEKKLHYQNLWTKLLGCQKNV  1581

Qy   891  EDWQKILMVRSLIVVSPHEDMRTWLKYIASLCGKSGRLALAHKTLVLLI--GVDPSRQLDHP  948
          :|||.:| |||:::|  :::|:.|:| ::.|||||||: || || | ||  | .:|.:..
Db  1582  DLWQRVLRVRSLVIVIKPKQDLQIWIKFANLCRKSGRMRLANKALNMLLEGGNDPS------ 1635

Qy   949  LPT---VHPQVTIYAYMKQNMWKSARKIDAFQHMQHFV------QTMQQQAQHAIAT   994
          ||    |  :|  :. ::| .:.:|  .: .: ::.:|      ||:.|.|:::.:|
Db  1636  LPNTVKAPPPVVYAQLKYIWATGAYKEALNHLIGFTSRIAHDLGLDPNNMIAQSVKLSSA  1695

Qy   995  EDQQHKQELHKQELHKLMARCFLKLGEWQLNLQGINESTIP-KVLQYYSAATEHDRSWYKAWHAW 1053
          ...  ::|:   ::|:|||||||::|||::.:| :|||:|. ||||| ::|:.|:|:|||||::|
Db  1696  STAPYVEEYTKLLARCFLKQGEWRIATQPMWRNTNPDAILGSYLLATHFDKNWYKAWHNW  1755

Qy  1054  AVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATATTTASTEGSNSESEA 1113
          |:.||| :|:| .:| |:|.
Db  1756  ALANFEVI------SMVQEETKLN--GGKNDDDDTAVN---NDNVRIDGSILGSGS      1801

Qy  1114  ESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVCQGFFRSISLSRGNNLQDTLRVLTLWFDY 1173
          .:.||.:   :|::       |  :|.|:||:.:|||.|||||.|. |||||.:|||:|::
Db  1802  LTINGNRYPLELIQR------HVVPAIKGFFHSISLLETSCLQDTLRLLTLLFNF       1850

Qy  1174  GHWPDVNEALVEGVKAIQDTWLQVIPLQIIARIDTPRRPLVGRLIHQLLLTDIGRYHPQALI 1233
          | .::.|||:.|.|.|: ::|:.:||||:|||:| |.::|:..||: |||:|:|:.|||||:
Db  1851  GGIKEVSQAMYBGFNLMKIENWLEVIPQLISRIHQPDPTVSNSLLSLLSDLGKAHPQALV  1910

Qy  1234  YPLTVASKSTTTARHNAANKILKNMCEHSNTIVQQAMMVSEELIRVAILWHEMWHEGLEE 1293
          ||||||.||:|:.|::||:.|:::.:..||:|||:|::|||||||||:||||:|::|||:
Db  1911  YPLTVAIKSESVSRQKAALSIIEKIRIHSPVLVNQAELVSHELIRVAVLWHELWYEGLED 1970

Qy  1294  ASRLYFGERNVAKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCKKYMKSGN 1353
          |.| :|.|.:|::.||.:||| |:.:.|.||||.|.|| :.||||::.|:||.:.|:|.:.
Db  1971  ARRQFFVEHNIERMFSTLEPIHKHLGNEPQTLSEVSFGRDLNDAYEWLNNYKKSKD      2030

Qy  1354  VKDLTQAWDLYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRI  1413
          :.:|.:|||:|:||:||::|:|||::|:||:||| ||:  .|||||||.|:|.:|.|||
Db  2031  INNLNQAWDIYNVFKRITRQIPQLQTLDLQHVSPQLLAATHDLELAVPGYFPDGKPTIRI 2090

Qy  1414  QSIAPSLQVITSKQRPRKLTIMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDP 1473
          :: .||.||:.|||||||::|:||.|:::|:.|||||:|||| :|||||||||||:|.|.
Db  2091  AKFEPLFSVISSKQRPRKFSIKGSDGKDYKVLKGHEDIRQDSLVMQLFGLVNTLLKNDS  2150

Qy  1474  TSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKLLNIEHRIMLRMAP  1533
          .. :|:|.|||||||||:..||::||||:.||:|.||:::|:.||:: ||:::|:|:|||
Db  2151  ECFKKHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHRDAKKIFINIEQWVMLQMAP 2210

Qy  1534  DYCHLTLMQKVBVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYI 1593
          || :|||:||:.||.:|::||.||:|.|:|||||.|||.|::||| |||||||||||:||
Db  2211  DYENLTLLQKIEVFTYALDNTKGDLLYKILWLKSRSSETWLERRTTYTRSLAVMSMTGYI 2270

Qy  1594  LGLGDRHPSNLMLDRLSGKTIHIDFGDCFEVAMTRKFPEKIPFRLTRMLTNAMEVTGLD  1653
          |||||||||||||||::||:|||||||||..|:.::||||:|||||||||.||||:||:
Db  2271  LGLGDRHPSNLMLDRITGKVIHIDFGDCFEAAILREKYPEKVPFRLTRMLTYAMEVSGIE 2330

Qy  1654  GNYRITCHTVMFVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQ 1713
          |.:||||.:||.|||::|:|:||:||||. ||| :|. :|.|.:.:: |:||||||||||
Db  2331  GSFRITCENVMRVLRDNKESLMAILEAFALDPLIHWGFDLPPQKLTBQT-----GI     2381

Qy  1714  SVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDF 1773
           :.:::.|||.|.|:||:  :|..:....|.:  :. |||:|| | :: |:.|.|||.|:
Db  2382  PLPLINPSEL---LRKGAITVERAANM--EAEQQNETRNARAMLVLRRITDKLTGNDI   2434

Qy  1774  SHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
          .:::.|||| || ||:|||||::||||.|||||||||
Db  2435  KRFNELDVPSQVDKLIQQATSIERLCQHYIGWCPFW 2470
```

RESULT 13
US-08-305-790B-3
; Sequence 3, Application US/08305790B
; Patent No. 6492106
GENERAL INFORMATION:
  APPLICANT: SABATINI, DAVID M.
  APPLICANT: ERDJUMENT-BROMAGE, HEDIYE
  APPLICANT: LUI, MARY
  APPLICANT: TEMPST, PAUL
  APPLICANT: SNYDER, SOLOMON H.
  TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
  TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION
  NUMBER OF SEQUENCES: 15
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: BANNER & ALLEGRETTI, LTD
    STREET: 1001 G STREET, N.W., 11TH FLOOR
    CITY: WASHINGTON
    STATE: D.C.
    COUNTRY: U.S.A.
    ZIP: 20001-4597
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/305,790B
    FILING DATE:
    CLASSIFICATION: 536
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/265,967
    FILING DATE: 27-JUN-1994
  ATTORNEY/AGENT INFORMATION:
    NAME: KAGAN, SARAH A.
    REGISTRATION NUMBER: 32,141
    REFERENCE/DOCKET NUMBER: 01107.47225
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 202-508-9100
    TELEFAX: 202-508-9299
    TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2470 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
  ORIGINAL SOURCE:
    ORGANISM: Saccharomyces cerevisiae
US-08-305-790B-3

```
Query Match         41.9%;  Score 3943;  DB 4;  Length 2470;
Best Local Similarity  44.2%;  Pred. No. 0;
Matches  812;  Conservative 360;  Mismatches 554;  Indels 110;  Gaps 26;

Qy    1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDDNPGVINNVLA  60
         |.||:   ::|:.:::|.|:.:|.:.|:.|:::.:|:|:|:|:|.|||.:... .::|
Db  718  LKFSTSSREKEETASLLCTLIRSSKDVAKPIEPLLNVLLPFKQ---DTSSTVASTALR 773
```

```
;  STREET:  1001 G STREET, N.W., 11TH FLOOR
;  CITY:  WASHINGTON
;  STATE:  D.C.
;  COUNTRY:  U.S.A.
;  ZIP:  20001-4597
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  Floppy disk
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/265,967C
;  FILING DATE:  27-JUN-1994
;  CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  KAGAN, SARAH A.
;  REGISTRATION NUMBER:  32,141
;  REFERENCE/DOCKET NUMBER:  01107.46363
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  202-508-9100
;  TELEFAX:  202-508-9299
;  TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  2:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH:  2470 amino acids
;  TYPE:  amino acid
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  protein
;  ORIGINAL SOURCE:
;  ORGANISM:  Saccharomyces cerevisiae
;  US-08-265-967C-2

Query Match            41.9%;    Score 3943;    DB 4;    Length 2470;
Best Local Similarity  44.2%;    Pred. No. 0;
Matches  812;    Conservative 360;    Mismatches 554;    Indels 110;    Gaps 26;

Qy      1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPVMEPILKALILLKDPDPDNPGVINNVLA   60
           ||| :  :||:: | || : :|  ||:| :|| ||: | :|| ::
Db    718  LKFSTSSREKEETASLLCTLIRSSKDVAKPYIEPLLNVLLPKFQ----DTSSTVASTALR  773

Qy     61  TIGELAQVSGLEMRKWVDELFIIIMDMLQDSLLAKRQVALWTLGQLVASTGYVVEPYRK  120
           ||||| | ||||| :| : :: :|||:|  :::||  ||: |||:||::||:
Db    774  TIGELSVVGGEDMKIYLKDLFPLLIKIFQDQSNSFKREAALKALGQLAASSGVVIDPLLD  833

Qy    121  YPTLEVILNFLKTEQNGTRRRFAIRVLGLLGALDPYEHKVANIGMIDQSRDAGAVSLSES  180
           | |:: :||| :| | |:|:||:||:| ||  |||:  ||:  |||:||  :::
Db    834  YPELLGILVNILKTENSQNIRRQVTVLTLGILGAIDPYRQK----------EREVTSTTDI  883

Qy    181  KSSQDSSDYSTSEMLVNMGNLPL-DEFYPAVSMVALMRIFRDQSLSHHHTMVVQATTFIF  239
           |:  : | ::|::  | ||| |  |  |  | ||: |||: |  : :||:||  |:||:
Db    884  STEQNAPPIDIA--LLMQGMSPSNDEYTTVVIHCLLKLIKDPSLSSYHHVAVIQAIMHIF  941

Qy    240  KSLGLKCVQPLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR  299
           ||:||||| ||:  | |: :| |: |  |||: ||||| :|| |:|:||:||  ||:||:
Db    942  QTLGLKCVSFLDQIIPTLLDVMRTCSQSLLEFYFQQLCSLIIIVRQHIRPFVDSIFQAIK  1001

Qy    300  EFWVMNTSIQSTILLIEQIVVALGGEFKLYLDPQLIPEMLRVFMEDNSPGRIVSIKLLAA  359
           : :  |:: ||||||:| | ::| |||:  ||| |::|| | :|::||| :||||||:||
Db   1002  DFSSV-AKLQITLVSVIEAISKALEGEFKRLVPLITLTLFIVILENDKSSDKVLSRRVLRL  1060

Qy    360  IQLFGANLDDYILHLLPPIVKLPDAPAPLPSRKAALETVDRLTBSLDFTYASRIIHPI  419
           ::|| |||:|:|:| :|: :  ::: ::|: | ::
Db   1061  LRKSFGPNLEGYSHLITPKIVQMAEFTSGNL--QRSAIIITGKJAKDVDLFEMSSRIVHSL  1118

Qy    420  VRTLDQ--SPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIV  477
           |||| |  : ||::|| :|::|:  | | ||:: |||:||:|||::|||||::
Db   1119  LRVLSSTTSDELSKVIMNTLSLLLIQMGTSPALFIPIVNEVLMKKHIQHTIYDDLCNRIL  1178

Qy    478  KGYTLADEEDPLIYQHRMLRSGCGDA-AASGPVETGP--MKKLHVSTINLQKAWGAARRV  535
           ||  ||:|::|||  ||:|:| | || ||||| | |   | |||||:|: ||| ||||
Db   1179  NNDVLPT--------KILBANTTDYKPAEQMEAADAGVAKLPINQSVLKSAWNSSQQR  1228

Qy    536  SKDDWLEWLRRLSLELLKDSSSPSLRSCWALLAQAYNPMARDLFNAAFVSCWSELNEDQQD  595
```

```
Matches 847; Conservative 333; Mismatches 514; Indels 169; Gaps 27;

Qy    1 LEHSGIGRIKEQSARMLGHLVSNAPRLLIRPYMEPILKALILKLKDFDPDNPGVINNVLA   60
          |||| : : :      |  :| | ::   ::  | ::: ::  :  |::  .:: :
Db  727 LKFSNMPKKKEESATLLCTLLNSSDEVAKPYIDPILDVILPKCQ---DASSAVASTALK  782

Qy   61 TTGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTIGQLVASTGYVVEPYRK  120
          :||| :  ::|:|:::|:|:: :| ::|||||::| || ||:| |||||:  :|| |::
Db  783 VLGELSVVGGKEMTRYLKELMPLIINTFQDQSNSFKRDAALTTIGQLAASSGYVVGPLLD  842

Qy  121 YPPLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHVKNIGMIIDQSRDASAVSLSES  180
          || ||  :| :||||:|  :|| :|:|| ||||||||||         :: : ::
Db  843 YPELLGGILINILKTENNPHIRRGTVRLLIGHLGALDPYKHR----------EIEVTSNS  890

Qy  181 KSSQDSSDYSTSEMLVNMGNLPL-DEFYPAVSMWALMRIFRDQSLSHHHMVVQAITFIF  239
          |||   .   : :: || | :|  ||:|| | |  |:  ::|:||| |  : :||  ||
Db  891 KSSVEQNAPSIDIALLMQGVSPSNDEYYPTVVIHNLMKIINDPSLSIHHTAIQAIMHIF  950

Qy  240 KSIGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR  299
          :.|||:||.|| |::| :: ::|.|     ||||||| |:|:||.|||| :.|| .|:
Db  951 QNIGLRCVSFLDQIIPGIILVMRSCPPSQLDFYFQQLGSLISIVKQHIRPHVEKIYGVIR 1010

Qy  300 EFWVMNTSIQSTILLIEQIVVALGGEBFKLYLPQLIPHMLRVFMDNSPGRIVSLKLLAA  359
          ||: :  | | |::: | | : |:   ||:  : | ::| :  | :|:||:|  |:||.
Db 1011 EFFPI-IKLQITIISVIEBSISKALEGEFKRFVPETLTFFLDILENDQSNKRIVPIRILKS 1069

Qy  360 IQLFGANLDDYLHILLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTYASRIIIPI  419
          :  ||.||:|| ||::|||| :  ::    .: |||  ::| |  :::.||:|||: .|
Db 1070 LVTFGPNLEDYSHLIMPIVVRMTEYSAGSL--KKISITIGRLAKNINLSEMSSRIVQAI 1127

Qy  420 VRTLDQSP-ELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVK  478
          ||  :     ||  | :||| |: |||.: :  :|:|| |: ::| || |||:: :.|
Db 1128 VRIIANNGDRELTKATMNTLSLLLQLGTDFVVFVPVINKALLRNRIQHSVYDLQLVNKLLN 1187

Qy  479 GYTLA------DEEEDPLIYQHRMLRSGGQDALASGPVETGPMKKLHVSTINLQKAWGAA  532
          ||:        :::  | ::::   :    :  : |:|||  |:::  :|| || |
Db 1188 NECLPTNIIFDKENEVP--------ERKNYEDEM--------QVTKLPVNQNILKNAWYCS 1232

Qy  533 RRVSKDDWLEWLRRLSLELLKDSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNED  592
          :: :|:||:|||:|||::|||| ||: |||: | |:||| |  || ||.|||.||  |
Db 1233 QQKTKEDWQEWIRRLSIQLLKESPSACLRSCSSLVSVYYPLAREHFNASFSSCWVELQTS 1292

Qy  593 QQDELIRSIELAL-TSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCR  651
          ::|:|::::  ||    |:  : : | || ||||| ||.|||:      :  ::|  :|
Db 1293 YQEDLIQALCKALSSSENPPEIYQMLNLVEFMEHDDK-PLPI----PIHTLGKYAQKCH 1347

Qy  652 AYAKALHYKELEFQKGPTPAILESLISINNKLQQPERAAGVLEYAMKHPGELEIQATWYE  711
          |.|||||||| || :   : | |||||||:|.|:: ||.|:| |||:||| | |:||||
Db 1348 AFAKALHYKEVEFLEEPKNSTIEALISINNQLRQTDSAIGILKHAQQH-NELQLKETWYE 1406

Qy  712 KLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAK  771
          || :|||| . |:::    : |  |: |:  |  ||||| :|    ::| | :::  :|
Db 1407 KLQRWEDALAAYNEKEAAGEBDSVEVMGKLRSLYALGEWEELSKLASERKWGTAKPEVKKA 1466

Qy  772 MARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLL  831
          ||  ||.||||.||:  ::|| :: .    |.|||:|:| || :: .|:  :  ||||||
Db 1467 MAPLAAGAAWGLEQWDEIAQYTSVMKSQSPDKEFYDAILCLHRNNFKKAEVHIFNARDLL 1526

Qy  832 DAELTAMAGESYSRAYGAMVSCHMLSLEEVIQYKLVP--ERRELIRQIWERLQQGCQR  888
          .:||:    :|  :|  :: : :::|: |:|:||| |   : |||  :|: |  ||:|:
Db 1527 VTELSALVNESTNRAINVVVRAQIIAELEEIIKYKKLPQNSDKRLITMRETWNTRLLGCQK 1586

Qy  889 IVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTVLLLGVDPSRQLDHP  948
          .::||::|| ||||:: | ||:.:::|:|:||:|||:|||.:| | .||  : :|||.|
Db 1587 NIDWQRILRVRSLVIKPKEDAQVRIKFANLCRKSGRMALAKKVLNTLL--EBTDDDDHP 1644

Qy  949 -LPTVHPQVTYAYMKNMWKSARKIDAFQHMQHF-----------VQTMQQCAQHA  991
           : ||.|.| |: | :||:::|:::| :  |:.            ..|:||.  ::
Db 1645 NTAKASPVVIAQYKYLWATGL-QDEALKQLINFTSRMAHDLGLDPNNMLAQSVPQQSKRV 1704

Qy  992 IATEDQQHKQELHKLMARCFLKLGEWGLNLQGINESTIPK-----------VLQYSAATEHD 1043
          ||:|| :::::|:|:  |||||.|||:  :|     :               ::||::|  |
Db 1705 -----PRHVEDYTKLLLARCFLKQGEWRVCLQ-----------PKWRLSNPDSIIGSYLLATHFD 1752
```

```
Qy 1580 YTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPEKIPFRL 1639
          |||||||||| ||||||||||||||||||| ||:| |||||||| |  |||| |:|||||
Db 2261 YTRSLAVMSMTGYILGLGDRHPSNLMLDRITGKVIIHIDFGDCFEAAILREKFPEKVPFRL 2320

Qy 1640 TRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGN 1699
          ||||: ||||:|::|::|||| :|||||| :| |:||:|||| :||||||  : |   :
Db 2321 TRMLITYAMEVSGIEGSFRITCENVMKVLRDNKGSLMAILEAFAFDPLLNWG-FDLPTK-- 2377

Qy 1700 KRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVP-ESIHSFIGDGLVKPEAL----- 1753
                              |::  ||:|:|::|  |:  :|:. |::: :  : :::
Db 2378 --------------------KIEEETGIQLPVMNANELLSNGAITEEEVQRVEN 2411

Qy 1754 ------NKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLLKKQATSHENLCQCYIGWC 1806
                 :|:||  :.||:||||||.|    ::::|||:|:| |:||||| |||||.|||||
Db 2412 EHKNAIRNARAMLVLKRITDKLTGNDIRRFNDLDVPBQVDKLIQQATSVENLCQHYIGWC 2471

Qy 1807 PFW 1809
          |||
Db 2472 PFW 2474
```

RESULT 11
US-08-305-790B-4
; Sequence 4, Application US/083057908
; Patent No. 6492106
; GENERAL INFORMATION:
; APPLICANT: SABATINI, DAVID M.
; APPLICANT: ERDJUMENT-BROMAGE, HEDIYE
; APPLICANT: LUI, MARY
; APPLICANT: TEMPST, PAUL
; APPLICANT: SNYDER, SOLOMON H.
; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
; TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & ALLEGRETTI, LTD
; STREET: 1001 G STREET, N.W., 11TH FLOOR
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,790B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,967
; FILING DATE: 27-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.47225
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 17430 BBMB UT
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
US-08-305-790B-4

Query Match   43.7%;   Score 4115.5;   DB 4;   Length 2474;
Best Local Similarity   45.5%;   Pred. No. 0;

CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & ALLEGRETTI, LTD
STREET: 1001 G STREET, N.W., 11TH FLOOR
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/265,967C
FILING DATE: 27-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.46363
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
US-08-265-967C-3

Query Match   43.7%;   Score 4115.5; DB 4;   Length 2474;
Best Local Similarity 45.5%; Pred. No. 0;
Matches 847; Conservative 333; Mismatches 514; Indels 169; Gaps 27;

```
Qy     1 LEEHSGIGRIKEQSARMLGHLVSNAPRLLRPYMEPILKALILLKDPDPDNPGVINNVLA     60
Db   727 LKFSNMPKKKEESATILCTILNSSDEVAKPYIDPILDVILPKCQ----DASSAVASTALK  782

Qy    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGLVASTGYVVEPYRK   120
Db   783 VLGELSVVGGKEMTRYLKELMPLIINTFQDQSNSFKRDAALTTLGQLAASSGYVVGPLLD  842

Qy   121 VPTLLEVLLNFLKTEQNGGITRREAIRVLGLLGALDPYKGH                    180
Db   843 YPELLGILINILKTENNPHIRRGTVRLIGILGALDPYKHR------------ELEVTSNS  890

Qy   181 KSSQDSDSYDYSTSEMLVNMGNLPL-DEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITPIF 239
Db   891 KSSVEQNAPSIDIALLMQGVSPSNDEYYPTVVIHNLMKILLNDPSLSIHHTAAIQAIMHIF  950

Qy   240 KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQOLGMLVSFVKSHIRPYMDEIVTLMR   299
Db   951 QNLGLRCVSFLDQIIPGTILVMRSCPPSOLDFYFQQLGSLISIVKQHIRPHVEKIYGVIR  1010

Qy   300 EFWVMNTSIQSTILILIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAA   359
Db  1011 EFFPI-IKLQITTIISVIESISKALEBGEFKRFVPBTLTFFLDILENDQSNKRIVPIRIIKS 1069

Qy   360 IQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLLTESLDFTDYASRIIHPI  419
Db  1070 LVTFGPNLEDYSHLIVPIVVRMTEYSAGSL--KKISI1TLGRLAKNINLSEMSSRIVQAL  1127

Qy   420 VRTLDQSP-ELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINEQRYDVLICRIVK   478
Db  1128 VRIINNGDREITKATMNTISLLLLQLGTDFVVFVPVINKALLRNRIQHSVTDQLVNKLLN  1187

Qy   479 GYTLA-----DEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLLHVSTNLQKAWGAA   532
Db  1188 NECLPTNIIFDKENEVP------ERKNYEDEM------QVTKLPLPVNQNILKNAWYCS  1232

Qy   533 RRVSKDDWLEWLRRISLSLLLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNED  592
Db  1233 QQKTKEDWQEWIRRISIQLLKESPSACLRSCSSLSVSVYYPLARELFNASFSSCWTELQTS 1292

Qy   593 QQDELIRSIELAL-TSQDIAEVTIQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCR  651
Db  1293 YQEDLIQALCKALSSSENPPEYQMLINLVEFMEHDK-PLPI---PIHTGKYAQKCH       1347

Qy   652 AYAKALHYKELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYE   711
Db  1348 APAKALHYKEVEFLEEPKNSTIEALLSINNQLHQTDSAIGILKHAQQH-NELQLKETWYE  1406

Qy   712 KLHEWEDALVAYDKKMDTNKDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAK   771
Db  1407 KLQRWEDALAAYNEKEAAGEDSVEVMWGKLRSLYALGEWERLSKLLASEKWGTAKPEVKKA 1466

Qy   772 MARMAAAAAWGLGWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLL   831
Db  1467 MAPLAAGAAGWGLEQWDEIAQYTSVMKSQSPDKEFYDAILCLHRNNFKKAEVHIFNARDLL 1526

Qy   832 DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVP--ERREIIRQIWERLGQCQR    888
Db  1527 VTELLSALVNRSYNRAYNVVVRAQIIAELEELIKYKKLPQNSDKRLTMRETWNTRLLGCQK 1586

Qy   889 IVEDWQKILMYRSLVVSPHEDMRTWLKYASLCGKSGRIALAHKTIVLLLGVDPSRQLDE?  948
Db  1587 NIDVWQRILRVRSLVIKPKEDAQVRIKFANLCRKSGRMALAKKVLNTLL--EETDDPDHP  1644

Qy   949 -LPTVHPQVTYAYMNMWKSARKIDAFQHMQHF-------------------VQTMQQQAQHA  991
Db  1645 NTAKASPPVVYAQLKYLWATGLQDEALKQLINFTSRMAHDLGLDPNNMIAQSVPQQSKRV  1704

Qy   992 IATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPK---------VLQYYSAATEHD 1043
Db  1705 -----PRHVEDYTKILLARCFLKQGEWRVCLQ-------PKWRLSNPDSILGSYLLATHFD 1752

Qy  1044 RSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATATTTAS 1103
Db  1753 NTWYKAWHNWALANFEVI-----------------SMLTSVSKKK               1780

Qy  1104 TEGSNSBEEARSTE--NSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNL  1160
Db  1781 QEGSDASSVTDINEFDNGMIGVNTFDAKEVHYSSNLIERHVIPAIKGFFHSISLSESSSL 1840

Qy  1161 QDTLRVLIWPDYGHWPDVNEALVEGVKAIQIDWLQVIPQLIARIDTPRPLVGRLIHQL   1220
Db  1841 QDALRLLTLWFTFGGIPEATQAMHEGFNLIQIGTWLEVLPQLISRIHQPNQIVSRSLLSL 1900

Qy  1221 LTDIGRYFPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVA  1280
Db  1901 LSDLGKAHPQALVYPLMVAILKSESLSRQKAALSIIEKMRIHSPVLVDQAELVSHELIRMA 1960

Qy  1281 ILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSPNQAYGRDLME 1340
Db  1961 VLWHRQWYGLGDDASRQFFGEHNTEKMFAAIL3PLYEMLKRGPETLREISFQNSFGRDLND 2020

Qy  1341 AQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYYSPKLLMCRDLELAV 1400
Db  2021 AYEWLMNYKKSKDVSNLNQAWDIYYNVFRKIGKQ-PQLQTLELQHVSPKLLSAHDLELAV 2080

Qy  1401 PGT-YDPNQPIIRIQSIAPSLQVITSKQRPRKLITLMGSNGHEFVFLLKGHEDLRQDERVM 1459
Db  2081 PGTRASGGKFIVKISKPEPVFSVISSKQRPRKFCIKGSDGKDYKYVLKGHEDIRQDSLVM  2140

Qy  1460 QLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKI 1519
Db  2141 QLFGLVNTLLQNDAECFRRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHREAKKI 2200

Qy  1520 LLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTPAGDDLAKLLWLKSPSSEVWFDRRTN 1579
Db  2201 PLNIEHWVMLQVAPDYDNLLLQKVEVFTYALNNTEGQDILYKVLWLKSRSSETWLERRTT  2260
```

```
Qy  1681  FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  1740
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  2421  FVYDPLLNWRLMDTNAKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  2480

Qy  1741  SPIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2481  SPIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  2540

Qy  1801  CYIGWCPFW  1809
          |||||||||
Db  2541  CYIGWCPFW  2549
```

RESULT 9
US-08-471-112A-4
; Sequence 4, Application US/08471112A
; Patent No. 6313264
; GENERAL INFORMATION:
; APPLICANT: Molnar-Kimber, Katherine L.
; APPLICANT: Failli, Amedeo F.
; APPLICANT: Caggiano, Thomas J.
; APPLICANT: Nakanishi, Koji
; APPLICANT: Chen, Yanqiu
; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,112A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/384,524
; FILING DATE: 13-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/312,023
; FILING DATE: 26-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,975
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Siekman, Michael T.
; REGISTRATION NUMBER: 36,276
; REFERENCE/DOCKET NUMBER: 01142.0058-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-471-112A-4

Query Match 49.8%; Score 4690; DB 4; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 894; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy   915  KYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAF   974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246  KYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAF   305

Qy   975  QHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ  1034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306  QHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ   365

Qy  1035  YYSAATEHDRSWYKAWEAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATT  1094
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366  YYSAATEHDRSWYKAWEAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATT   425

Qy  1095  AATATTTASTEGNSESEABSTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL  1154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426  AATATTTASTEGNSESEABSTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL   485

Qy  1155  SRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPPRPLVG  1214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486  SRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPPRPLVG   545

Qy  1215  RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSE  1274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546  RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSE   605

Qy  1275  ELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY  1334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606  ELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY   665

Qy  1335  GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVSPKLLMCR  1394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666  GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVSPKLLMCR   725

Qy  1395  DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQ  1454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726  DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQ   785

Qy  1455  DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR  1514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786  DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR   845

Qy  1515  EKKKILLNIEHRIMLRMAPDYDHLTMQKVEVEHAVNNTAGDDLAKLLWLKSPSSEVWF  1574
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846  EKKKILLNIEHRIMLRMAPDYDHLTMQKVEVEHAVNNTAGDDLAKLLWLKSPSSEVWF   905

Qy  1575  DRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEK  1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906  DRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEK   965

Qy  1635  IPFRLTRMLTNAMEVTGLDGNYRITCHTMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT  1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966  IPFRLTRMLTNAMEVTGLDGNYRITCHTMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT  1025

Qy  1695  NTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN  1754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1026  NTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN  1085

Qy  1755  KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW  1809
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1086  KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW  1140
```

RESULT 10
US-08-265-967C-3
; Sequence 3, Application US/08265967C
; Patent No. 6476200
; GENERAL INFORMATION:
; APPLICANT: SABATINI, DAVID M.
; APPLICANT: ERDJUMENT-BROMAGE, HEDIYE
; APPLICANT: LUI, MARY
; APPLICANT: TEMPST, PAUL
; APPLICANT: SNYDER, SOLOMON H.
; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
; TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION
; NUMBER OF SEQUENCES: 14

```
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/305,790B
      FILING DATE:
      CLASSIFICATION:  536
  PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/265,967
      FILING DATE:  27-JUN-1994
  ATTORNEY/AGENT INFORMATION:
      NAME:  KAGAN, SARAH A.
      REGISTRATION NUMBER:  32,141
      REFERENCE/DOCKET NUMBER:  01107.47225
  TELECOMMUNICATION INFORMATION:
      TELEPHONE:  202-508-9100
      TELEFAX:  202-508-9299
      TELEX:  197430 BBMB UT
  INFORMATION FOR SEQ ID NO:  2:
      SEQUENCE CHARACTERISTICS:
      LENGTH:  2549 amino acids
      TYPE:  amino acid
      TOPOLOGY:  linear
      MOLECULE TYPE:  protein
      HYPOTHETICAL:  YES
      ANTI-SENSE:  NO
      ORIGINAL SOURCE:
      ORGANISM:  Rattus rattus
      TISSUE TYPE:  pheochromocytoma
      CELL TYPE:  PC12
  US-08-305-790B-2
```

Query Match   99.1%;   Score 9325;   DB 4;   Length 2549;
Best Local Similarity 99.0%;   Pred. No. 0;
Matches 1791; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Db   741 LEHSGIRIKEQSARMLGHLVSNAPRLRPYMBPILKALILKLKDPDPDNPGVINNVLA 800
Qy   61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
              |||||||||||||||||||||||:|:||||||||||||||||||||||
Db   801 TIGELAQVSGLEMRKWVDELFVIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 860

Qy   121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861 YPTLLEVLLNFLKTEQNQGTRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 980

Qy   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy   301 FWVMNTSIQSTILLIBQIVVALGGEFKLYLPQLIPHMLRVTMHDNSPGRIVSIKLLAAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1041 FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSQGRIVSIKLLAAI 1100

Qy   361 QLFGANLDDYLHLLLPPIYKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1101 QLFGANLDDYLHLLLPPIYKLFDAPEVPLPSRKAALETVDRLTBSLDFTDYASRIIHPIV 1160

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLLICRIVKGY 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLLCRIVKGY 1220

Qy   481 TLADEEDDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1221 TLADEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy   601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy   661 ELEPQKGPTPAILESLISINNKLQQPEAAAGVLEYAMCHFGELEIQATWYEKLHEWEDAL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1401 ELEFQKGPTPAILESLISINNKLQQPEAASGVLEYAMCHFGELEIQATWTEKLHEWEDAL 1460

Qy   721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy   781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDILLDAELTAMAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDILLDAELTAMAG 1580

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 1640

Qy   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHFLPTVHPQVTYAY 1700

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy 1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy 1081 SGANITNAITIAATTAATTATTTASTEGGSNSESRAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
              |||||||||||||||||||||||||||||||:|||:|||||||||||||||::|
Db 1821 SGANITNAITTTAATTAASAAAATSTEGSNSESEAESNESSPTPSPLQKKVTEDLSKTLLLY 1880

Qy 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy 1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHQALLYPLTVASKSTTARHNAANKILKNMCE 1260
              ||||||||||||||||||||||||:||||||||||||||||||||||||||
Db 1941 QLIARIDTPRPLVGRLIHQLLDIGRYHPQALLYPLTVASKSTTARHNAANKILKNMCE 2000

Qy 1261 HSNTLVQQAMMVSEELIRVALIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2001 HSNTLVQQAMMVSEELIRVALIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy 1321 GPQTLKETSFNQAYGRDLMEAQEWCKRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2061 GPQTLKETSFNQAYGRDLMEAQEWCKRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy 1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2121 LELQYVSPKLLMCRDLELAVPGTYDPNQTIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy 1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy 1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEFEHAVNNTAGDDLA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDELTLMQKVEFEHAVNNTAGDDLA 2300

Qy 1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMLDRLSGKILHIDFGD 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy 1621 CFEVAMTREKFPEKIPPFRLTRMLINAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2361 CFEVAMTREKFPEKIPFFRLTRMLINAMEVTGLDRNYRTTCHTVMEVLREHKDSVMAVLEA 2420

Qy 1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2421 FVYDPLLNWRLMDTNAKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy 1741 SFIGDGLVKPEALNKKAIQINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2481 SFIGDGLVKPEALNKKAIQINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy 1801 CYIGWCPFW 1809
              |||||||||
Db 2541 CYIGWCPFW 2549

RESULT 8
US-08-305-790B-2
; Sequence 2, Application US/08305790B
; Patent No. 6492106
; GENERAL INFORMATION:
;   APPLICANT: SABATINI, DAVID M.
;   APPLICANT: ERDJUMENT-BROMAGE, HEDIYE
;   APPLICANT: LUI, MARY
;   APPLICANT: TEMPST, PAUL
;   APPLICANT: SNYDER, SOLOMON H.
;   TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
;   TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION
;   NUMBER OF SEQUENCES: 15
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: BANNER & ALLEGRETTI, LTD
;   STREET: 1001 G STREET, N.W., 11TH FLOOR
;   CITY: WASHINGTON
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20001-4597
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible

Qy   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy   601 IELALTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLJGERAAKCRAYAKALHYK 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 IELALTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLJGERAAKCRAYAKALHYK 1400

Qy   661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYKLHEWEDAL 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYKLHEWEDAL 1460

Qy   721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy   781 WGLGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 WGLGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 1640

Qy   901 SLVVSPHEDMRTWLKYASLCGKSGRIALAHKTIVLLGVDPSRQLDHPLPTVHPQVTYAY 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 SLVVSPHEDMRTWLKYASLCGKSGRIALAHKTIVLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy   1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy   1081 SGANITNATTAATTAATATTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1821 SGANITNATTAATTAATATTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880

Qy   1141 TVPAVGQFFRSISLSRGNNLQDTLRVLTIWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881 TVPAVGQFFRSISLSRGNNLQDTLRVLTIWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy   1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKLIKNMCE 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARINAANKLIKNMCE 2000

Qy   1261 HSNTLVQQAMMVSBELIRVAILWHEMWHGLEEASRLYFGSRNVKGMFEVLEPLHAMMER 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2001 HSNTLVQQAMMVSBELIRVAILWHEMWHGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy   1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy   1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy   1441 EFVFLIKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2181 EFVFLIKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSN.MLDRLSGKILHIDFGD 2360

Qy   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680

Db   2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy   1681 FVYDPLLNWRLMDTNYKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2421 FVYDPLLNWRLMDTNYKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy   1801 CYIGWCPFW 1809
           ||| |||
Db   2541 CYIGWYPFW 2549

RESULT 7
US-08-265-967C-1
; Sequence 1, Application US/08265967C
; Patent No. 6476200
; GENERAL INFORMATION:
; APPLICANT: SABATINI, DAVID M.
; APPLICANT: ERDJUMENT-BROMAGE, HEDIYE
; APPLICANT: LUI, MARY
; APPLICANT: TEMPST, PAUL
; APPLICANT: SNYDER, SOLOMON H.
; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
; TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: BANNER & ALLEGRETTI, LTD
;   STREET: 1001 G STREET, N.W., 11TH FLOOR
;   CITY: WASHINGTON
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20001-4597
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/265,967C
;   FILING DATE: 27-JUN-1994
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: KAGAN, SARAH A.
;   REGISTRATION NUMBER: 32,141
;   REFERENCE/DOCKET NUMBER: 01107.46363
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-508-9100
;   TELEFAX: 202-508-9299
;   TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2549 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: YES
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;   ORGANISM: Rattus rattus
;   TISSUE TYPE: pheochromocytoma
;   CELL TYPE: PC12
;
US-08-265-967C-1

Query Match   99.1%; Score 9325; DB 4; Length 2549;
Best Local Similarity   99.0%; Pred. No. 0;
Matches 1791; Conservative   5; Mismatches   13; Indels   0; Gaps   0;

Qy   1 LEHSGIGRIKEQSARMLGHLVSNAPRLLRPYMEPLLKALILKLKDPDPDNPGYINNVLA   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,112A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/384,524
FILING DATE: 13-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/312,023
FILING DATE: 26-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,975
FILING DATE: 08-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Siekman, Michael T.
REGISTRATION NUMBER: 36,276
REFERENCE/DOCKET NUMBER: 01142.0058-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2549 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-471-112A-3

Query Match   99.9%; Score 9402; DB 4; Length 2549;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 1808; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKLKDPDPDNPGVINNVLA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKLKDPDPDNPGVINNVLA   800

Qy   61 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK   860

Qy   121 YPTTLEVLLNFLKTEQNQGTRRRAIRVLGLLGALDPYKHKVNIGMLDQSRDASAVSLSES   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861 YPTTLEVLLNFLKTEQNQGTRRRAIRVLGLLGALDPYKHKVNIGMLDQSRDASAVSLSES   920

Qy   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK   980

Qy   241 SLGLKCVQFLPQVMPFFLNVIRVCDGAIREFLFQQLGMIVSFVKSHIRPYMDEIVTLMRE   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981 SLGLKCVQFLPQVMPFFLNVIRVCDGAIREFLFQQLGMIVSFVKSHIRPYMDEIVTLMRE 1040

Qy   301 FVWMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 FVWMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMEDNSPGRIVSIKLLAAI 1100

Qy   361 QLFGANLDDYLHLLLPPIVLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIHHPIV   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLLLPPIVLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIHHPIV 1160

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy   481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Db   1821 SGANITNATTAATTAATATTTASTEGSNSBESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141 TVPAVQGFFRSISLSRGNNLQDTLRVLLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881 TVPAVQGFFRSISLSRGNNLQDTLRVLLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy   1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKLKNMCE   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKLKNMCE   2000

Qy   1261 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2001 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy   1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy   1381 LELQVVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPKLTLMGSNGH 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2121 LELQVVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPKLTLMGSNGH 2180

Qy   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRVAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy   1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPRSIH 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPBSIH 2480

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2481 SFIGDGLVKPEALNKKAIQIINRVRDKITGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy   1801 CYIGWCPFW 1809
          |||||||||
Db   2541 CYIGWCPFW 2549

RESULT 6
US-08-471-112A-3
; Sequence 3, Application US/08471112A
; Patent No. 6313264
; GENERAL INFORMATION:
; APPLICANT: Molnar-Kimber, Katherine L.
; APPLICANT: Failli, Amedeo F.
; APPLICANT: Caggiano, Thomas J.
; APPLICANT: Nakanishi, Koji
; APPLICANT: Chen, Yangiu
; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;   ADDRESSEE: Dunner, L.L.P.
;   STREET: 1300 I Street, N.W.
;   CITY: Washington
;   STATE: DC
;   COUNTRY: USA
;   ZIP: 20005-3315
; COMPUTER READABLE FORM:

```
Qy      1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPIIKALILKLKDKPDPDNPGVINNVLA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPIIKALILKLKDKPDPDNPGVINNVLA 800

Qy     61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASGYYVVEPYRK 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASGYYVVEPYRK 860

Qy    121 YPTLLEVLNFLKTEQNGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    861 YPTLLEVLNFLKTEQNGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy    181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy    241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy    301 FWVMNTSIQSTILLLEQIVALGGEFKLYLPQLIPHMLRVIMHDNSPGRIVSIKLLAAI 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 FWVMNTSIQSTILLLIEQIVVALGGEFKLYLPQLIPHMLRVFMEDNSPGRIVSIKLLAAI 1100

Qy    361 QLFGANLDDYLHLLPPIVKLFDAPEAPLPSRKAALETVDRLIESLDFTDYASRIIHPIV 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASR=IHPIV 1160

Qy    421 RTLDQSPELRSTAMDTLSSAVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy    481 TLADEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 TLADEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy    541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELLRS 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELLRS 1340

Qy    601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy    661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy    721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKVARMAAAAA 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKVARMAAAAA 1520

Qy    781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy    841 ESYSRAYGAMVSCHMLSELEVIQYKLVPERREIIRQIWWERLQGGCQRIVEDWQKILMVR 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ESYSRAYGAMVSCHMLSELEVIQYKLVPERREIIRQIWWERLQGGCQRIVEDWQKILMVR 1640

Qy    901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy    961 MKNMWKSARKIDAFQEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKIGEWQLN 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 MKNMWKSARKIDAFQEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKIGEWQLN 1760

Qy   1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy   1081 SGANTINATTNATTAATATTTASTEGSNSESEAASTENSPTPSLQKKVTEDLSKTLLMY 1140
```

```
Db   1261 HSNTLVQQAMVSSELIRVAIIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321 GPCTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GPCTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380

Qy   1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440

Qy   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500

Qy   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMWGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMWGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620

Qy   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680

Qy   1681 FVYDPLLNRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 FVYDPLLNRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800

Qy   1801 CYIGWCPFW 1809
            |||||||||
Db   1801 CYIGWCPFW 1809
```

RESULT 5
PCT-US95-06722-12
; Sequence 12, Application PC/TUS9506722
; GENERAL INFORMATION:
;   APPLICANT:
;   TITLE OF INVENTION: Immunosuppressant Target Proteins
;   NUMBER OF SEQUENCES: 25
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII (text)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US95/06722
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/250,795
;   FILING DATE: 27-MAY-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/250,795
;   FILING DATE: 20-DEC-1994
; INFORMATION FOR SEQ ID NO: 12:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2549 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
PCT-US95-06722-12

Query Match   100.0%;   Score 9413;   DB 5;   Length 2549;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1809;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Db  1681 FVYDPLLNWRLMTNTKGNKRSRRTDSYSAGQSVEILDGVELGEPAHKKTGTVPESIH 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800

Qy  1801 CYIGWCPFW 1809
         |||||||||
Db  1801 CYIGWCPFW 1809
```

RESULT 4
US-09-012-399A-12
; Sequence 12, Application US/09012399A
; Patent No. 6500152
; GENERAL INFORMATION:
; APPLICANT: Berlin, Vivian
; APPLICANT: Chiu, Maria Isabel
; APPLICANT: Cottarel, Guillaume
; APPLICANT: Damagnez, Veronique
; TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: FOLEY, HOAG & ELIOT LLP
;   STREET: One Post Office Square
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109-2170
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/012,399A
;   FILING DATE:
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/360,144
;   FILING DATE: 20-DEC-1994
; ATTORNEY/AGENT INFORMATION:
;   NAME: Vincent, Matthew P.
;   REGISTRATION NUMBER: 36,709
;   REFERENCE/DOCKET NUMBER: APV-036.02
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-832-1000
;   TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1809 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-012-399A-12

Query Match   100.0%; Score 9413; DB 4; Length 1809;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1809; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 LEHSGIRIKEQSARMLGHLVSNAPRJIRPYMEPILKALIIKLKDPDPDPNPGVINNVLA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 LEHSGIRIKEQSARMLGHLVSNAPRJIRPYMEPILKALIIKLKDPDPDPNPGVINNVLA 60

Qy   61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120

Qy  121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180

Qy  181 KSSQDSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 KSSQDSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIPRDQSLSHHHTMVVQAITFIFK 240

Qy  241 SIGLKCVQFLPQVMPFTLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 SIGLKCVQFLPQVMPFTLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300

Qy  301 FWVMNTSIQSTILLLIEQIVVALGGEFKLYLPQLIPHMLRVFMDNSPGRIVSIKLLAAI 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 FWVMNTSIQSTILLLIEQIVVALGGEFKLYLPQLIPHMLRVFMDNSPGRIVSIKLLAAI 360

Qy  361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420

Qy  421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480

Qy  481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540

Qy  541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600

Qy  601 IELALTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLGERAAKCRAYAKALHYK 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 IELALTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLGERAAKCRAYAKALHYK 660

Qy  661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAI 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAI 720

Qy  721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGLHQQCCEKWTLVNDETQAKMARMAAAAA 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGLHQQCCEKWTLVNDETQAKMARMAAAAA 780

Qy  781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVIALHQDLFSLAQQCIDKARDLLDAELTANAG 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVIALHQDLFSLAQQCIDKARDLLDAELTAMAG 840

Qy  841 ESYSRAYGAMVSCHMLSBLEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 ESYSRAYGAMVSCHMLSBLEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 900

Qy  901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVILLGVDPSRQLDHPLPTVHPQVTYAY 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVILLGVDPSRQLDHPLPTVHPQVTYAY 960

Qy  961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020

Qy 1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKLRHA 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKLRHA 1080

Qy 1081 SGANINATTAATTAATATTTASTEGSNSESEARESTENSPTPSPLQKKVTEDLSKTLLMY 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 SGANINATTAATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLNY 1140

Qy 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLITLMFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLITLMFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200

Qy 1201 QLIARIDTPRPLVGRLIHQLLTIDIGRYHPQALLYPLTVASKSTTARHNAANKIIKNMCE 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 QLIARIDTPRPLVGRLIHQLLTIDIGRYHPQALLYPLTVASKSTTARHNAANKIIKNMCE 1260

Qy 1261 HSNTLVCQAMMVSEELIRVALIWHEMWHHGLEEASRIYFGERNVKGMFEVLEPLHAMMER 1320
```

```
Db   1801 CYIGWCPFW 1809

RESULT 3
US-09-012-504A-12
; Sequence 12, Application US/09012504A
; Patent No. 6464974
; GENERAL INFORMATION:
; APPLICANT: Berlin, V.
; APPLICANT: Chiu, I.
; APPLICANT: Cottarel, G.
; APPLICANT: Damagnez, V.
; TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
; FILE REFERENCE: APBI-P05-036
; CURRENT APPLICATION NUMBER: US/09/012,504A
; CURRENT FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 08/360,144
; PRIOR FILING DATE: 1994-12-20
; PRIOR APPLICATION NUMBER: 08/250,795
; PRIOR FILING DATE: 1994-05-27
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
; LENGTH: 1809
; TYPE: PRT
; ORGANISM: Mammalian
US-09-012-504A-12

Query Match             100.0%;   Score 9413;   DB 4;   Length 1809;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1809;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA 60

Qy    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120

Qy   121 YPTTLEVLLNFLKTEQNGGTRREAIRVLGLLGALDPYKHVNIGMIDQSRDASAVSLSES 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 YPTTLEVLLNFLKTEQNGGTRREAIRVLGLLGALDPYKHVNIGMIDQSRDASAVSLSES 180

Qy   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 240

Qy   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQOLGMLVSFVKSHIRPYMDEIVTLMRE 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQOLGMLVSFVKSHIRPYMDEIVTLMRE 300

Qy   301 FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPFMLRVFMHDNSPGRIVSIKLLAAI 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPFMLRVFMHDNSPGRIVSIKLLAAI 360

Qy   361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKGAALETVDRLTESLDFTDYASRIIHPIV 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480

Qy   481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540

Qy   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600

Qy   601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660

Qy   661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720

Qy   721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780

Qy   781 WGLGGWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 WGLGGWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 900

Qy   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 960

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLH 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020

Qy  1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080

Qy  1081 SGANITNATTAATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 SGANITNATTAATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140

Qy  1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEAIVEGVKAIQIDTWLQVIP 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEAIVEGVKAIQIDTWLQVIP 1200

Qy  1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKMMCE 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260

Qy  1261 HSNTLVQQAMMVSEELIRVAILWHEMWEEGLEEASRLYPGERNVKGMFEVLEPLHAMMER 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 HSNTLVQQAMMVSEELIRVAILWHEMWEEGLEEASRLYPGERNVKGMFEVLEPLHAMMER 1320

Qy  1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380

Qy  1381 LELQVVSPKLLMCRDLELAVPGTYDPNOPIIRIQSIAPSLQVTSKQRPRKLLTMGSNGH 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 LELQVVSPKLLMCRDLELAVPGTYDPNOPIIRIQSIAPSLQVTSKQRPRKLLTMGSNGH 1440

Qy  1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500

Qy  1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRVAPDYDHLTMQKVEVPEHAVNNTAGDDLA 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTMQKVEVPEHAVNNTAGDDLA 1560

Qy  1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKLLHIDFGD 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKLLHIDFGD 1620

Qy  1621 CFEVAMTREKFPEKIPFRLRMLINAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CFEVAMTREKFPEKIPFRLRMLINAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680

Qy  1681 FVYDPLLNWRLMDTNTKGNKRSRIPTDSYSAGQSVEILDGVELGEPAHKTGTTVPESIH 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    SOFTWARE:  PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/360,144A
    FILING DATE:  20-DEC-1994
    CLASSIFICATION:  435
 ATTORNEY/AGENT INFORMATION:
    NAME:  Vincent, Matthew P.
    REGISTRATION NUMBER:  36,709
    REFERENCE/DOCKET NUMBER:  APV-036.02
 TELECOMMUNICATION INFORMATION:
    TELEPHONE:  617-832-1000
    TELEFAX:  617-832-7000
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
    LENGTH:  1809 amino acids
    TYPE:  amino acid
    TOPOLOGY:  linear
 MOLECULE TYPE:  protein
US-08-360-144A-12
```

```
Query Match             100.0%;  Score 9413;  DB 3;  Length 1809;
Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1809;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy     1 LEHSGIGRIKQSARMLGHIVSNAPRLIRPYMEPILKALIILCKDPPDPDENPGVINNVLA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 LEHSGIGRIKQSARMLGHIVSNAPRLIRPYMEPILKALIILCKDPPDPDENPGVINNVLA 60

Qy    61 TIGHLAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVVEEYRK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TIGHLAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVVEEYRK 120

Qy   121 YPTILEVLLNFLKTEQNGTRREAIRVLGLLGALDPYHKVNIGMIDQSRDASAVSLSES 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 YPTILEVLLNFLKTEQNGTRREAIRVLGLLGALDPYHKVNIGMIDQSRDASAVSLSES 180

Qy   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240

Qy   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEVTLMRE 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEVTLMRE 300

Qy   301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360

Qy   361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHEQRYDVLICRIVKGY 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHEQRYDVLICRIVKGY 480

Qy   481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQXAWGAAARRVSKDDW 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQXAWGAAARRVSKDDW 540

Qy   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600

Qy   601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660

Qy   661 ELEFQKGPTPAILESLIISNNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ELEFQKGPTPAILESLIISNNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720

Qy   721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780

Qy   781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 900

Qy   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 960

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCPLKLGEWQLN 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCPLKLGEWQLN 1020

Qy  1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080

Qy  1081 SGANINATTNATTAATTASTEGSNESBAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 SGANINATTNATTAATTATTTASTEGSNSSSEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140

Qy  1141 TVPAVQGFFRSISLSRGNNLQDTLRVILMFDYGHWPDVNEALVBGVKAIQIDTWLQVIP 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TVPAVQGFFRSISLSRGNNLQDTLRVILMFDYGHWPDVNEALVBGVKAIQIDTWLQVIP 1200

Qy  1201 QLIARIDTPRPIVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 QLIARIDTPRPIVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260

Qy  1261 HSNTLVQQAMVSEELIRVALIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 HSNTLVQQAMVSEELIRVAIIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPIHAMMER 1320

Qy  1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTS 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTS 1380

Qy  1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTIMGSNGH 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTIMGSNGH 1440

Qy  1441 BFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 BFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500

Qy  1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560

Qy  1561 KLLWIKSPSSEVWFDRRTNYTRSLAVMSMVGYILGIGDRHPSNLMLDRLSGKILHIDFGD 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 KLLWIKSPSSEVWFDRRTNYTRSLAVMSMVGYILGIGDRHPSNLMLDRLSGKILHIDFGD 1620

Qy  1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680

Qy  1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILLDGVEILGEPAHKKTGTTVPESIH 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILLDGVEILGEPAHKKTGTTVPESIH 1740

Qy  1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800

Qy  1801 CYIGWCPFW 1809
         |||||||||
```

```
Db 1081 SGANITNATTAATTAATATTASTEGSNSESEARSTENSPTTPSPLQKKVTEDLSKTLLMY 1140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200

Qy 1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTPARHNAANKILKNMCE 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260

Qy 1261 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLRPLHAMMER 1320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 HSNTLVQQAMVVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLRPLHAMMER 1320

Qy 1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380

Qy 1381 LELQVSPKLLMCRDLELAVPGTYDFNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 LELQVSPKLLMCRDLELAVPGTYDFNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440

Qy 1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRRNLSIQRYAVIPLSTNSGLIGWV 1500
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRRNLSIQRYAVIPLSTNSGLIGWV 1500

Qy 1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEYFEHAVNNTAGDDLA 1560
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEYFEHAVNNTAGDDLA 1560

Qy 1561 KLLWLKSPSSEWFDRRTNYTRSLAVMSMVGYIGLGDRHPSNLMLDRLSGKIIHIDFGD 1620
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 KLLWLKSPSSEWFDRRTNYTRSLAVMSMVGYIGLGDRHPSNLMLDRLSGKIIHIDFGD 1620

Qy 1621 CPEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 CPEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680

Qy 1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740

Qy 1741 SFIGDGLVKPEALNKKAIQINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 SFIGDGLVKPEALNKKAIQINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800

Qy 1801 CYIGWCPFW 1809
        |||||||||
Db 1801 CYIGWCPFW 1809
```

```
Db                                                              
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPTILKALJILKLKDPDPDNPGVINNVLA 60

Qy   61 TIGELAQVSGLEMRKNVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TIGELAQVSGLEMRKNVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPIRK 120

Qy  121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180

Qy  181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240

Qy  241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300

Qy  301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSKLLAAI 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSKLLAAI 360

Qy  361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLLDFTDYASRIIHPIV 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLLDFTDYASRIIHPIV 420

Qy  421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480

Qy  481 TLADEEEDPLIYCHRMLRSGQGDALASGPVETGPMKKLIVSTINLQKAWGAARRVSKDDW 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TLADEEEDPLIYCHRMLRSGQGDALASGPVETGPMKKLIVSTINLQKAWGAARRVSKDDW 540

Qy  541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNEMARDLFNAAFVSCWSELNEDQQDELIRS 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNEMARDLFNAAFVSCWSELNEDQQDELIRS 600

Qy  601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660

Qy  661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720

Qy  721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGGLHQQCCEKWTLVNDETQAKMARMARAAAA 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGGLHQQCCEKWTLVNDETQAKMARMARAAAA 780

Qy  781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLAILHQDLFSLAQQCIDKARDLLDAELTAMAG 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLAILHQDLFSLAQQCIDKARDLLDAELTAMAG 840

Qy  841 ESYSRAYGAMVSCHMLSELEBVIQYKLVPERREIIRQIWWERLQGCCRIVEDWQKILMVR 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 ESYSRAYGAMVSCHMLSELEBVIQYKLVPERREIIRQIWWERLQGCCRIVEDWQKILMVR 900

Qy  901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVILLGVDPSRQLDHPLPTVHPQVTYAY 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVILLGVDPSRQLDHPLPTVHPQVTYAY 960

Qy  961 MKNMWKSARKIDAFQHMQHFVQTMQQQACHAIATEDQQHKQELHKLNARCFLKLGEWQLN 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 MKNMWKSARKIDAFQHMQHFVQTMQQQACHAIATEDQQHKQELHKLNARCFLKLGEWQLN 1020

Qy 1021 LQGINESTTPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 LQGINESTTPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080

Qy 1081 SGANITNATTAATTAATATTASTEGSNSESEARSTENSPTTPSPLQKKVTEDLSKTLLMY 1140
```

```
                GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      March  2, 2004, 19:59:58 ; Search time 30 Seconds
                                         (without alignments)
                                         3113.046 Million cell updates/sec

Title:       US-09-517-491-12
Perfect score: 9413
Sequence:    1 LEHSGIGRIKEQSARMLGHI.........KQATSHENLCQCYIGWCPFW 1809

Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5

Searched:    389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:    389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued Patents AA:*
              1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
              2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
              3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
              4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
              5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
              6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

              Pred. No. is the number of results predicted by chance to have a
              score greater than or equal to the score of the result being printed,
              and is derived by analysis of the total score distribution.

                                      SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9413	100.0	1809	3	US-09-012-515A-12	Sequence 12, Appl
2	9413	100.0	1809	3	US-08-360-144A-12	Sequence 12, Appl
3	9413	100.0	1809	4	US-09-012-504A-12	Sequence 12, Appl
4	9413	100.0	1809	4	US-09-012-399A-12	Sequence 12, Appl
5	9413	100.0	2549	5	PCT-US95-06722-12	Sequence 12, Appl
6	9402	99.9	2549	4	US-08-471-112A-3	Sequence 3, Appli
7	9325	99.1	2549	4	US-08-265-967C-1	Sequence 1, Appli
8	9325	99.1	2549	4	US-08-305-790B-2	Sequence 2, Appli
9	4690	49.8	1140	4	US-08-471-112A-4	Sequence 4, Appli
10	4115.5	43.7	2474	4	US-08-265-967C-3	Sequence 3, Appli
11	4115.5	43.7	2474	4	US-08-305-790B-4	Sequence 4, Appli
12	3943	41.9	2470	4	US-08-265-967C-2	Sequence 2, Appli
13	3943	41.9	2470	4	US-08-305-790B-3	Sequence 3, Appli
14	1735	18.4	562	3	US-09-012-515A-14	Sequence 14, Appl
15	1735	18.4	562	3	US-08-360-144A-14	Sequence 14, Appl
16	1735	18.4	562	4	US-09-012-504A-14	Sequence 14, Appl
17	1735	18.4	562	4	US-09-012-399A-14	Sequence 14, Appl
18	1735	18.4	562	5	PCT-US95-06722-14	Sequence 14, Appl
19	848	9.0	162	3	US-09-012-515A-2	Sequence 2, Appli
20	848	9.0	162	3	US-08-360-144A-2	Sequence 2, Appli
21	848	9.0	162	4	US-09-012-504A-2	Sequence 2, Appli
22	848	9.0	162	4	US-09-012-399A-2	Sequence 2, Appli
23	848	9.0	162	5	PCT-US95-06722-2	Sequence 2, Appli
24	751.5	8.0	2644	4	US-09-029-047C-2	Sequence 2, Appli
25	725.5	7.7	2386	4	US-08-609-049A-26	Sequence 4, Appli
26	714	7.6	184	3	US-09-170-996-26	Sequence 26, Appl
27	714	7.6	184	3	US-09-170-996-26	Sequence 26, Appl

28	693	7.4	2930	4	US-09-417-822-2	Sequence 2, Appli
29	634	6.7	3056	1	US-09-360-416-2	Sequence 8, Appli
30	631	6.7	3056	2	US-08-508-836A-8	Sequence 3, Appli
31	631	6.7	3056	2	US-08-629-001A-3	Sequence 3, Appli
32	631	6.7	3056	3	US-08-874-266-2	Sequence 3, Appli
33	631	6.7	3056	3	US-08-642-274D-3	Sequence 3, Appli
34	631	6.7	3056	3	US-08-952-127-3	Sequence 3, Appli
35	631	6.7	3056	4	US-08-952-014C-3	Sequence 2, Appli
36	631	6.7	3056	4	US-08-984-090-2	Sequence 2, Appli
37	630.5	6.7	3057	4	US-09-360-416-3	Sequence 3, Appli
38	630	6.7	230	1	US-09-481-620A-111	Sequence 111, App
39	628.5	6.7	3066	3	US-08-952-127-12	Sequence 12, Appli
40	611.5	6.5	1708	1	US-08-493-092-2	Sequence 2, Appli
41	611.5	6.5	1708	1	US-08-508-836A-2	Sequence 2, Appli
42	570	6.1	2368	1	US-08-198-446B-15	Sequence 15, Appli
43	570	6.1	2368	2	US-08-870-693-15	Sequence 15, Appli
44	544	5.8	129	1	US-09-481-620A-79	Sequence 79, Appli
45	542	5.8	100	2	US-09-963-601-2	Sequence 2, Appli

```
                                    ALIGNMENTS

RESULT 1
US-09-012-515A-12
; Sequence 12, Application US/09012515A
; Patent No. 6127521
;  GENERAL INFORMATION:
;   APPLICANT: Chiu, Maria Isabel
;   APPLICANT: Cottarel, Guillaume
;   APPLICANT: Damagnez, Veronique
;   TITLE OF INVENTION:  IMMUNOSUPPRESSANT TARGET PROTEINS
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;    ADDRESSEE: FOLEY, HOAG & ELIOT LLP
;    STREET:  One Post Office Square
;    CITY:  Boston
;    STATE:  MA
;    COUNTRY:  USA
;    ZIP:  02109-2170
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE:  Floppy disk
;    COMPUTER:  IBM PC Compatible
;    OPERATING SYSTEM:  PC-DOS/MS-DOS
;    SOFTWARE:  PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/09/012,515A
;    FILING DATE:
;    CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 08/360,144
;    FILING DATE:  20-DEC-1994
;  ATTORNEY/AGENT INFORMATION:
;    NAME:  Vincent, Matthew P.
;    REGISTRATION NUMBER:  36,709
;    REFERENCE/DOCKET NUMBER:  APV-036.02
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  617-832-1000
;    TELEFAX:  617-832-7000
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;    LENGTH:  1809 amino acids
;    TYPE:  amino acid
;    TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-012-515A-12

Query Match             100.0%;  Score 9413;  DB 3;  Length 1809;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1809;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDDPNPGVINNVLA 60
```

```
Qy   1166  VLTLWFDYGHMPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIG 1225
           :||||:|:.||    :|:.  |    :|: ||||||:|||||||||||| ||||  ::|
Db    414  LLTLWFKYGYQBEVCSAIKEGYGSVNIDTWLEVIPQLIARIHAPDQKVRRLIHDLLAEVG 473

Qy   1226  RYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHE 1285
           :||||||||||||||||||    |   | :|||: : |||::||:.| :||||||||||
Db    474  KEHPQALIYPLTVASKSQSFNRQTAASSVMDKLRAHSAVLVEQALLVSQELIRVAILWHE 533

Qy   1286  MWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWC 1345
           :||:||||||||    :|: :   :||  :| |: ||:|:||: | |||:| ||:|||
Db    534  IWMENLEEASRLFFIERNFEAMFNILDELQERLLDRGPETLREVAFVQAYGRELQEAGGWC 593

Qy   1346  RKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPG 1402
           :||.:: |: :  |:|| |: :|: |: |:|   :| |||||||||   || ||||||
Db    594  KKYKQTGDMTFLSQAWQIYHEMFKKITAQVGTITTLELQYVSPKLLAARDLELAVPG 650
```

Search completed: March 2, 2004, 20:03:21
Job time : 73 secs

Qy   1704 TRTDSYSA-GQSVEILD------GVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKA 1757
                 :|: : || || || |:| || | |:|| |:|
Db   1471 SSVQTISVIGSMTNVAEGSVGSLQLSQLSVQRRATT-----QQRVNANIEKDTSQTRKP 1525

Qy   1758 IQIINRVRDKLJTGRDF------SHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
                 :| : || | ||||||| ||:|||:||||||:|||||||||:||
Db   1526 KSVVKRIKRKLRGLEFPQSQEGKSSDGFTVEEQVSRLIEEATSNENLCVHFLGWCPFW 1583

RESULT 14
Q8T9I6
ID   Q8T9I6   PRELIMINARY;   PRT;   455 AA.
AC   Q8T9I6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   GM10438p.
GN   TOR OR CG5092.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunco J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY066275, AAL39420.1; -.
DR   FlyBase; FBgn0021796; Tor.
DR   GO; GO:0030307; P:positive regulation of cell growth; NAS.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3KC; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE   455 AA; 51951 MW; E7314609A35C1C92 CRC64;

Query Match   17.7%; Score 1667; DB 5; Length 455;
Best Local Similarity   68.8%; Pred. No.1e-113;
Matches 321; Conservative 51; Mismatches 69; Indels 26; Gaps   4;

Qy   1356 DLTQAWDLIYHVFRRISKQLPQLTSLELQIYSPKLLMCRDLELAVPGTYDPNQPIIRIQS 1415
                |||:|||||:||:||: || |||||||:|| || |||: |:||:||| ||| |:|||
Db   2 DLDRAWDIYHVHFQK1SRQLPQLTSLELPYVSPKLMTCKDLELAVPGSYNPGQELIRISI 61

Qy   1416 IAPSLQVITSKQRPRKLJTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTS 1475
                |:|||||||||||||| |||: |::|||||:||||||||||||||:|||||||||||:
Db   62 IKTNLQVITSKQRPRKLJCIRGSNGKDYMYLLKGHEDLRQDERVMQLFSLVNTLLLDDPDT 121

Qy   1476 LRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILNIEHRIMLRMAPDY 1535
                |:||:||||:|| ||||||||||||||||:|||||||||||:|:|||:|||:|:|:|||
Db   122 FRRNLAIQRYAVIPLSTNSGLIGWVPHCDTLHTLIRDYRDKKKKVPLNQEHRTMLNFAPDY 181

Qy   1536 DHLTLMQKVEFEHAVNNTAGDDLAKLLMLWLKSPSSEVWFDRRCNYTRSLAVMSMVGYILG 1595
                |||||||||||:||:||||||||||||||||||||:|||:||||:|||||||||||||||
Db   182 DHLTLMQKVEFEHALGQTQGDDLAKLLMLWLKSPSSELWFERRNNYTRSLAVMSMVGYILG 241

Qy   1596 LGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLJTRMLTNAMEVTGLDGN 1655
                ||||||||||:||||||||||||||||||||||||||||||||||||||| |:||:||:|
Db   242 LGDRHPSNLMLDRMSGKILHIDFGDCFEVAMTREKFPEKIPFRLJTRMLIKAMEVTGIEGT 301

Qy   1656 YRITCHTVMEVLREHKDSVMVLRAFVYDPLLNWRLMDINTKGNKSRRTJTDSYS----- 1710
                :|:|||:|||||:||||| |: |:|: |||| |: :| | || |:
Db   302 YRRTCESVMVLRRNKDSLMAVLEAFVVDPLLNWRLLDVDKKGN------DAVAGAGAP 354

Qy   1711 -----AGQSVEILDGVELGEPAHKK--TGTTVPESIHSFIGDGLVKPEALNKKAIQIIN 1762

Db   355 GGRGGSGWQDSLSNSVEDSLPVAKSKPYDPTLQQGGLENNVAD------ETNSKASQVIK 408

Qy   1763 RVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
                | :|:: || :|||||||||:|||:|||||:|||||||||||||||
Db   409 RVKCKLTGTDFQTEKSVNEQSQVELLIQQATNNENLCQCYIGWCPFW 455

RESULT 15
Q9HFM9
ID   Q9HFM9   PRELIMINARY;   PRT;   651 AA.
AC   Q9HFM9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   TOR2 protein (Target of rapamycin) (Fragment).
GN   TOR2.
OS   Glomus mosseae.
OC   Eukaryota; Fungi; Glomeromycota; Glomeromycetes; Glomerales;
OC   Glomeraceae; Glomus.
OX   NCBI_TaxID=27381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BEG12;
RA   Requena N., Mann P., Franken P.;
RT   "A homologue of the cell cycle check point TOR2 from yeast cells exits
RT   in the arbuscular mycorrhizal fungus Glomus mosseae.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ276633; CAC08177.1; -.
DR   HSSP; P42345; 1FAP.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF02259; FAT; 1.
FT   NON_TER   1   1
FT   NON_TER   651   651
SQ   SEQUENCE   651 AA; 75596 MW; 420C6C1B4AAF4F15 CRC64;

Query Match   17.2%; Score 1621; DB 3; Length 651;
Best Local Similarity   46.2%; Pred. No. 4.5e-110;
Matches 331; Conservative 117; Mismatches 197; Indels 72; Gaps   8;

Qy   691 GVLEYAMK--HFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALG 748
                :||:| :: : ::||: | ||||||::|||:   : : |:||||||::|: |||
Db   1 GILTHAQQNPNLQDIQLKENWYEKLNWEDALTAYEKRTDDPDNDNEATLGVMRCLHSLG 60

Qy   749 BWGQLHQQCCEKWTLVNDETQAKMARMARAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRA 808
                |||:||| :|: || || ||:| | ||| |||| | ||: : :|||:|:| |: |
Db   61 DMDGLSELVQERKWPRAETEMRKSIATYAAGAWSLGQWGLMEDYIKNIKSDNYYRPFYQA 120

Qy   809 VIALHQDLFSLAQQCIDKARDLLDABLTAMAGESYSRAYGAMVSCEMLSELBEVIQYLJV 868
                |:|||:|| : | : ::|:|::| | ||:| ||||:|| ||||:|||: |||||:||::|
Db   121 IIALHKNQYSBAVRYIDKTRDLLDTELTALVGESYNRAYNVVVRVQMLABLBEIITTYK-- 178

Qy   869 PERRRIIRQIWWERLQGGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLAL 928
                :| |:::| | || : :|::| | | || :: |:||| | |:| | |||::|
Db   179 ------QYAQGVPRNVDVWQQLKVRALVISPKENMEMWIKFANLCRKSDRLEF 226

Qy   929 AHKTLVLLLGVDPSRQLDHP--LPTVHPQVTYAYMKNMWKSARKIDAFGHMQHFVQTMQQ 986
                | ::||||:|| |::| : : |:| ||| | ||:| ||:|:| |:|||:|| | |||
Db   227 AEKTLKLLLVGDQKLELIHPSLLARXSPQVVYAHLKWWASXVREEALEFLRDPTTRM-- 284

Qy   987 QAQHAIATBDQQHKQELHKLMARCFLKLJGEWQLNLQ-GINESTIPKVLQYYSAATEHDRS 1045
                   :| ::| :||||| |: |: : || :|:| | | | |::| :| || |
Db   285 ------SNNDRVGNNGLLARCYLKKGEWQKVLQDDWDRETVPDILQSYLLATQYDKD 335

Qy   1046 WYKAWHAVAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATATTASTE 1105
                ||||||| |: : : ||:: |: |||:|:| ||||| | ||||
Db   336   WYKAWHAWALANLEV1SFV---------ERLKSAEN-----------NTPADQTITSVTE 375

Qy   1106 GSNSESEARSTENSPFPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLLSRGNNLQDTLR 1165
                  ::| | | : | || | | | | || || ||||||:| ||||||||||||
Db   376 ------------------------------EISGDIQQEVLRYVYRPFKGFFRSIALSRGNCLQDTLR 413

Qy   1711 -----AGQSVEILDGVELGEPAHKK--TGTTVPESIHSFIGDGLVKPEALNKKAIQIIN 1762

```
RESULT 13
O76222
ID   O76222;        PRELIMINARY;       PRT;   1583 AA.
AC   O76222;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Phosphatidylinositol 3 kinase (Fragment).
GN   PI3K.
OS   Trypanosoma brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5691;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AnTat1;
RX   MEDLINE=98418771; PubMed=9747975;
RA   Bringaud F., Vedrenne C., Cuvillier A., Parzy D., Baltz D., Tetaud E.,
RA   Pays E., Venegas J., Merlin G., Baltz T.;
RT   "Conserved organization of genes in trypanosomatids.";
RL   Mol. Biochem. Parasitol. 94:249-264(1998).
DR   EMBL; AF031925; AAC32769.1; -.
DR   PIR; T14176; T14176.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Kinase.
FT   NON_TER       1      1
SQ   SEQUENCE   1583 AA;  178357 MW;  0F3CE4A79E9323C5 CRC64;

Query Match             21.6%;  Score 2036.5;  DB 5;  Length 1583;
Best Local Similarity   32.1%;  Pred. No. 5.5e-140;
Matches  539;  Conservative 297;  Mismatches 643;  Indels 199;  Gaps 45;

Qy    232 VQAITFIF--KSLGLKC-VQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIR 288
              |:::  ||  .|: ||  |:  :: |:|:|   ||| | ::| |: |::|:||:.
Db      5 VQVIVGIFSLEVLSPGCLVSFYGEVISSIQKQARV-QVRKREILVHLISLVRILKGHLR  63

Qy    289 PYMDEIVTLMREF-WVMNTSIQSTILLIEQIVVALGGEFKLYLFQLI-PHMLRVFMHDN 346
            | :|: || ||  :| : ||   : |||: ||: || :::| ||:  :: |::|| :
Db     64 PHLKEITSTVDSFISVTDLSVLRQVLALLKELCCSLRRRFRPYMSSLIGPIIVLVEENVE 123

Qy    347 SPGRIVSIKLLAAIQLFGANLLDDYLHLLLPPIVKLFDPAPAPLPSRKAALETVDRLTESL 406
           |||:|   |:||:   |    ::  :| | :   |:||: | |  ||  ||::|||:||:|
Db    124 ETSEIV---LDFFSAMGSLLEDHLHTVLPVVCNIIVDTSVPSRCRIVAVKTLICTKRL 179

Qy    407 -DFTDYASRIIHPIVRTLDQSP--------ELRSTAMDTLSSLVFQLGKKYQIFIPMV- 455
            :: : |:   || :|  ||||          ||:| |:|| ||:| |:||| |:|:|
Db    180 PDLCFHASRCVHCLCRVLRESDGDGDGVGDBEGRLGCSAMEALCTLAGSLGKNFENFVLVL 239

Qy    456 NKVLVRHRINHQRYDVLICRIVKG-YTLADEEBDPLIYQHRMLRSGQGDALASGPVBTG- 513
             :  :|  :   || : ||  : |  :  ::|  ||: |: | :  |    :: ||||
Db    240 PNVADRYGETSGEY----CRFCHDIYEAIDGKRAPEVSSNGVGKAGGGGA-PSLPFTAGT 294

Qy    514 ---PMKKLHVSTINLQKAWGAARRVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAY 570
              |::|| :::: |:: :  | :::|:||| |:| ||:| ::::  |::|| || |: |
Db    295 SASPLKDRADAYASLRFHLRKRDQASDEDWNHWLPQLAVNLLRSSSSPSHGRALTLAELH 354

Qy    571 NPMARDLFNAAFVSCWSELNEDQQDELIRSIELALTSQDI-AEVTQTLLNLAEFMEH--- 626
            || || | || |:| || |:| ||:|||  |: |:  |:  |: :  :| || ||
Db    355 EPFARQMLHSAFAACYADMDEHTQREVIGLLTEVIRGLRVPSEVMQELLNLSEHMERQGI 414

Qy    627 -SDKGPLPLRDDNGIV----LLGERAAKCRAYAKALHYKELEFQK----------- 666
            : .   ||  : | |:      ||:| || |  |:|:|  ||  :
Db    415 RLSAGGKASIKHSSNFCPFDRQVLMBSSANCNLYSKALHYABIEFLETVREYERS-ILRGC 474

Qy    667 -GPTPA--------ILESLISINNKLQQPEAAAGVLEYAM------------KHFGELE--IQA 707
            ||  |         |:|:|| :   | ||:|: :  :||            ||| || |  ||
Db    475 PKPLPVEDWQNLIKLCEKSIYFCNLLGQRESANGIIKFIRQNLPLLGKKVTELSQMMDA 534

Qy    708 TWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTIVNDE 767
            :||:|| ||:|| ||  ||  |  :||  |:||:|:|||||  :|  | ||  |:| |:
Db    535 HLFDKLRWWSQSLQAYERRLQQEPNKVSNMVGLLRALDALGEYPRVLEMWRQFSKRVSKR 594

Qy    768 TQAKMARMAAAAAWGLGQWDSMEBYTCMIPRDTHDGA---FYRAVLAIHQDLFSLAQQCI 824
            : |: ::: :||   ||:  |::|:|  |||  | |     ||:|| :||:|| |:|||:
Db    595 EVSKLASMGAHAAANLLRRWDDMEHITSFMSDEDYTGTTALFYKATLAARKKKFREAEKLI 654

Qy    825 DKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEVIQYKLVPERREIIRQIWWERLQ 884
            |  | :| |: | :  :::   || : :| |:||||:| :| :|  | :  |:|  :::
Db    655 DMCRKRVDSKLSALVAESYDRAYDLFVGIQQLSELELAMATSDPQSAMHRQLWERRLS 714

Qy    885 GCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCCKSGRLALAHKTLVLLLGVDP-SR 943
            || ::|:|||:|:|||:||:||| :|  :||   |  | ||:|||: :|: | |::| :|
Db    715 --VMAYEGWPGTLANHTLLVPPSSEIDMWLRFVSLSRAHGQGSVSTEVLRELGNQSIES 772

Qy    944 QLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQQAHAIATEDQQHKQEL 1003
            ||: || : |   :: :   :| |  |: |::|||||| | |: ||  |:||:| ||:|
Db    773 AIENGIPT--PAVAMGSFQHLYETNQRDSAIARLQLYVSKVEGGSAQHVSRRED----- 825

Qy   1004 HKLMARCFLKLGEWQLNLQGIN--ESTIPKVLQYYSAATEHDRSWYKAHHAWAVNNFEAV 1061
             |:: : | : ||  |: |:     |:||:: : || ::  | |:|:| :: ||||: |
Db    826 ---MAVCHAKLAEWLVHQKKAHRTEDELQKIFHHLRRATELDKSNGSIWRTLARVHREA- 881

Qy   1062 LHYKHQNQARDEKKKLRHASGANITNATTAATTAAAATATTTASTEGSNSESEAESTENSPT 1121
             :     |:   |:  |:  || :   :|  :: : | | : :      ::
Db    882 ---------------------AIKPADGSDSSGASG------AAG---- 896

Qy   1122 PSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLILWFDYDYGHWP---- 1177
            :  | : :  :   |:|: :
Db    897 ---------------HIMEALSAYLRSVSLS--EELEDALGFLSLWFMVG--PLLAV 934

Qy   1178 DVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLT 1237
            :|  :   .|  | |:  |||: ||: || |  |:| |:| |: | ||  |:||| | :
Db    935 QVGSTLKEEIBEVNPTVWLKVIPQIIARISSSNGTVADSVYNLLVLVARRHPQAILYSLN 994

Qy   1238 VA-----SKSTTTARH-NAANKILKNMCE-HSN--TLVQQAMMVSEELIRVAILWHEMW 1287
            ||      | : : |  :||:|||:| |  ||    |:|:|: |:|:|||::|  || :
Db    995 VAHSSYQGKGTADGVEPLKGSHRVLARIAEHIQNGKAMVEDSALVCRELVRCAVLWPELW 1054

Qy   1288 HEGLERASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRK 1347
            |:|| :|::|::|   ::| |  : | |: :|  | :|:: |:: |  ||: :|:  |:|
Db   1055 FNELGRALYQWERQRSAENLLLAMGPLLEQLKR-PETMAEAQFAAELRQPLENACCHVER 1113

Qy   1348 YMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYYSPKLLM-CRDLELAVPGTY-- 1404
            | : |:::| |: ||: | |:  :|:: |  :|: :  |:|||:  |||:|   | |
Db   1114 AVSSRHEQFMEEARRIRFVSIERRIREQISGMSSLAIQLVSPKLHQNGRNLSLVVPGQYRE 1173

Qy   1405 DPNQPIIRIQSIAPSLQVTTSKQPRKGTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGL 1464
            | :  || : :: |:| :  ||| |  |:|:|:|   | | :||:|||  || |:||
Db   1174 DGNYPL--IASFQNVLKVLNSKQPRRIYINGSNGEIYKFLLKGHEDLRLDERVMQLLGF 1231

Qy   1465 VNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIE 1524
            ||:|  |:| ::|:||   | ::||| :    :| | :  :|:|: :|||  |:|| |::
Db   1232 VNTILEKHSVPKRHDCLIQTYSVTPLSDNAGLVGWVDHCDTLNKIIEDYRVNPR-CIRME 1290

Qy   1525 HRIMLRMAPDYDHLTLMQKVEVFEEAVNNTAGDDLAKLLWLKSPSSEVWFDRPTNYTRSL 1584
            | :::|  |:  |: :|: :  ::|   || || :||  |:| |:||||::| ::| | |
Db   1291 LDIMRSMCDNLIYLTAIQRVEPFEFPLEBTEGVDLVRSFWVKAPSAETWLRRTTYVVCSL 1350

Qy   1585 AVMSMVGYILGLGDRHPSNLMLDRLSGKTLHIDFGDCFEVAMTREKFPEKIPFRLTRMLT 1644
            |:|:|| :|||||||||:||: |:|: |:  |||||||  |:  | |  ::|| ||| |:
Db   1351 ATMSMVGHILGLGDRHPSNLMIHAFSGRVVHIDFGDCFEVAQQRSIHPEKVPFRLTRMLV 1410

Qy   1645 NAMEVTGLDGNYRITCHTVMBVLREHKDSVMAVLEAFVYDPLLN-WRLMDTNTKGNKRSR 1703
            :| || ||::||  |: |: |:||:|  |: | |:|::|| ::  |:| |||:|| |:
Db   1411 KAMEMGGIEGLFRHGCHTVMNVLREBGGSLLALLEAFVHDPLVSWWRDEAEGFSGNGQAA 1470
```

```
Qy   948 ------------------------------------------------------------  947

Db  1606 QHVALPTSDVLSIASHSVAKPTLQMIANASTHQGQSIEKETPLSIVGQQCDDAKRNAQRD 1665

Qy   948 ------------------------------------------------------------  947

Db  1666 LFNALIPPLNHVYPHVLYAYCKHLWVSSNENVLBKLVAFTHLIRLCNDVSIQ-----KTB 1720

Qy   996 DQQHKQELHKLMARCFIKLGBWQLNL----------------------QGINESTIPKVLQYYS 1037

Db  1721 P--------KLVCRMLIRLGAWYSDLLTPLDHKSCSMIGTISPSEGYVNIPVQDTVTSCCS 1772

Qy  1038 A----------ATEH----------------DRSWYK-------AWHAWAVM 1056

Db  1773 AIWNSIVNNCDADASVCSTTYEATEHSAERLELSSASTIDKAVAKLVDNILLAYDELSNV 1832

Qy  1057 NFEAVLHYKHQNQA----------------RDEK-----KKLRHASGANINATTA-- 1091

Db  1833 SLBECNQYNTVGMSRIGLIALEWFQVAADRDETSNHIWKCISHTTYA-IVNSFAAYIDKL 1891

Qy  1092 --------------ATTAATATTTASTEGSNSESEABSTENSPTP--SPLQKKVTED----L 1133

Db  1892 NDRPGVDGGLQIVTRSSVDATSDAVGSTSRALSQSVQDSDRPLVCPEHLGFVQNRQVLL 1951

Qy  1134 SKTLLMYTVPAVQGFFRSISLSRG-NNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQI 1192

Db  1952 EKAKLEYLVESIIAHFKCIETGRSIEALPTTIRLMTLWFRYGSYDEVTEIINGLAAVPI 2011

Qy  1193 DTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAAN 1252

Db  2012 NIWLDVIPQLIARLHSPQHKIRLLVHQLLVIGTEHPQALIYPLVVASKSTILNRRVESL 2071

Qy  1253 KILKNMCEHSNTIVQQAMMVSEELIRVAILWHEMWHEGLEEASRLIVFGE-RNVKGMFEVL 1311

Db  2072 SIVDQIRRTNEAMILESLINKELQRVAILWPEIAYDLITQASTAFCEBPRDLVKFPYA 2131

Qy  1312 EPLHAMMERGPQTLKETSFNQAYGRDLMEAQBWCRKYMKSGNVKDLTQAWDLYHVFRRI 1371

Db  2132 CQLQLLASRTPETQSESVFIQMFSSQLQAAWDSCRQFPVTNDLVCLNIVMKFYKEIYDRI 2191

Qy  1372 SKQLPQLTSLELQVYSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRK 1431

Db  2192 HSELPFLNTIDLLQTSPLLYNISNTNLCVPGSYHPARAITT!VKFYQKVFVIRSKQRPRK 2251

Qy  1432 LTLMGSNGHBFVFLLKGHBDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLS 1491

Db  2252 VGLVASDGBYYAFLLKGHBDLRQDERVMQLFTLINNLLMNDSYCSSRGLMVMKYPITPLS 2311

Qy  1492 TNSGLIGWVPHCDTLHALIRDYRBKKKILLNIEHRIMLRMAPDYDHITLMQKVEVFBHAV 1551

Db  2312 QNSGLLYWVPGCDTIISLIKEFRQLKRIHLBTEMQQIKYRAPAFMNMALHYKVDAFKYML 2371

Qy  1552 NNTAGDDLAKLLMLKSFSSEVWPDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMDRLSG 1611

Db  2372 TTSDAIDLCRSLMNKASSASBDWLLKRMNYTRSIAVSSIVGYIIGLGDRHPANLMIERTTG 2431

Qy  1612 KILHIDFGDCFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLRFHK 1671

Db  2432 MVLHIDYGDCFEVAMHRQLPEKVPFRLTSIMWKAFESCGTESSFRMTGEATMSILRSNK 2491

Qy  1672 DSVMAVLEAFVYDPLLNWRLMD-TNTKGNKRSRSRTRDSYSAGQSVEILLDGVELGEPAHKK 1730

Db  2492 DSLVSVLETFIYDPLMSMKVLSGENTMGGK----------ENIDG------NTEF 2530

Qy  1731 TGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSH-DDTLDVPTQVELLI 1789

Db  2531 YGTRRP--------VQESYLEGNERNPKAVKIVQRIYDKLITGVDIKYIDRSLSVSEQIQGSLI 2584

Qy  1790 KQATSHENLCQCYIGWCPFW 1809

Db  2585 HSATNAENLCQLYIGWCPYW 2604
```

```
Qy   281 SFVKSHIRPYMDEIVTLMREFWVMNTSIQSTILLIEQIVVAL----------GGEFKLYLPQ 333

Db   534 SSIGHHAHSYVSDIFDIIEVLWA-NQQLIELILALLEEVSIVLDLSLGNYGSQFFLKFVG 592

Qy   334 LIPHMLRVFMHDNSPGRIVSIKLLAAIQLFGANLDYLHLILPPIVKLFDAPEAPLPSRK 393

Db   593 LFT-VLKSSQEISDSAAKLLIRACRSVVALSLFSNHIHLIVDLLCSGIRKQKLGKELRF 651

Qy   394 AALEIVDRLTESLDFTDYASRIIHPIVATL--------DQSPELRSTAMDTLSSL 440

Db   652 ALLDTLSRLTICTNLSRFGASIIHPILEVIMEADPCTSVEHNKNANPEKGLNPMNIVSYK 711

Qy   441 VF----------QLGKKYQIFIPMVNKVLV-----------------RHRINHQRYDVLIC 474

Db   712 AFEVLHFICAQFGVSVTLYIPVITBCLLKVGFTSGLLDSIFSLLLRREEVDPVTFYRLIC 771

Qy   475 RIVK----------GYT------LADEH-----EDPLIYQHRMLRSGQQD- 503

Db   772 EWKKKQAKKFSEYTSLASLLAMYEFPVTQDTLLYMESISNKVPGDDKDAMILVVKDINFD 831

Qy   504 ----------------ALASGPVBTGPMKKLHVSTINLQKAWGAARRV 535

Db   832 DIGSVINYLHTSSTWFGDDMRPLPAGAIGVHSLSIQENDMLNKHVTI--PQQIWRADTCK 889

Qy   536 SKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSEL---NED 592

Db   890 KNEEWRHWLSNLSLTLLQCSPLKCLRACHKLATSYPRVARDLFNYSPLACLSETYPSNNE 949

Qy   593 QQDELIRSIELALTSQDI-AEVTQTLLNLABFMEHSDKGPLPLRDDNGIVLLGERAAKCR 651

Db   950 HRNYIVESIRSVFLSQNCPLDVLQILLDLABFWENRSREFSTTFSHN---FLGEVABRCH 1006

Qy   652 AYAKALHYKELEFQKGPTPAILESLISINNKLQQPBAAAGVLEYAMCHF---------- 700

Db  1007 AYTRSLRYKETEYAQHPIEAV-GRLIAINYSLGYEETAIGTLRVECKRLNPISTNLIYDI 1065

Qy   701 ------------GELEIQAT----------------------- 708

Db  1066 LLRCLYTGLDDVFTMMQREGGLEPGAAVNCSAESTLERIINKVVSKIVTKALTEAKKISN 1125

Qy   709 -------------------------------------------- 708

Db  1126 LNQSIPQQLVSYGSIGGIFVEGRPVMSGPEGSVLQVDYNEIGIYLSALAARRMLKSKLTYQ 1185

Qy   709 ------------WYEKLHEWEDALVAYDKKMD--------- 728

Db  1186 YPTFLMYEENLFQSVLKTLRYKAQGHMQDLYGPSFKWFENVCLWDKALSGYNRQIDQLKL 1245

Qy   729 ---TNKDDPBLML------GRMRCLEALGEWGQLHQQCCEKWTLVNDETQAK------- 771

Db  1246 SLKSNTSBARTLPPMLVNLIKDKIRCLFALADYSAVQBEADQLLNYLSSESKEBAGLITTV 1305

Qy   772 ------MARMAAAAAWGLGQWD--SMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQ 821

Db  1306 KHTLIKGFQADIATYAAWSLLEQDQEDVEIWLRVLPEDNIDGTIIRATSYIKANRMLZAH 1365

Qy   822 QCIDKARDLLDAELTAMAGSYSRAYGAMVSCHMLSBLEEVIQYKL---------- 867

Db  1366 KLLCKLRLRIDPDLIGLSABSYTRAYNLCILLQTTTELQNIISYKLALNGLKLDAPSTLI 1425

Qy   868 --------------------VPERREIIRQIWWBERLQGCCRIVEBDWQKILMVRSLV 903

Db  1426 KBGGNPYENIFETTLISFNPPGFADTEKERILRLMWSRKLISMIRWDCBSWEALLRVRRLI 1485

Qy   904 VSPHEDMRTWLKYASLCGKSGRLALAHKT---------LVLLLGV---- 939

Db  1486 IFPIEDQBVWIRFSVLCLQSGRVQLARKTLE*TPINSSSVIHIKVSKSRGIYTLLGLTSFG 1545

Qy   940 DPSRQLDH-------------------------------------- 947

Db  1546 TPRRYLKHGSGDVSTGGSVTPIGSALGAEMGVVQDLSAVGSSLAKYLLKDGVVKDAFTAQ 1605
```

```
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY052068; AAK93492.1; -.
DR   FlyBase; FBgn0021796; Tor.
DR   GO; GO:0030307; P:positive regulation of cell growth; NAS.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008940; Prenyl_trans.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE   760 AA;  86888 MW;  A081D76114080BE0 CRC64;

Query Match             28.1%;  Score 2646.5;  DB 5;  Length 760;
Best Local Similarity   63.8%;  Pred. No. 2.1e-185;
  Matches  521;  Conservative  92;  Mismatches 133;  Indels  71;  Gaps   9;

Qy     1007 MARCFLKLGEWQLNLQ-GINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYK 1065
            ||||:|:|:| ::|    :||||:: |:  :||||||
Db        1 MARCYLRMATWQNKLQDSIRPDAIQGALECFEKATSYDPNWYKAWHLWAYMNFKVV---Q 57

Qy     1066 HQNQARDEKKKLRLHASGANITNATTAATTAATATTTASTEGSNSESEAESTENSPTTPSPL 1125
                  |:|::                           :|:
Db       58 AQKSALLDKQQ----PPGASM-------------GWTMGGSGLLSDL-- 85

Qy     1126 QKKVTEDLSKTLLMTVPAVQGFFRSISLSRGNNLQDTLRVLTLMFDYGHWPDVNEALVE 1185
                  :  :                            ::|:|:|::|||||:
Db       86 -------------MIIQRYAVPAVQGFFRSISLIKGNSLQDTLRILITLMFDYGNHAEVYEALLS 136

Qy     1186 GVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTIDIGRYHPQALIYPLTVASKSTTT 1245
            |:|:|:|
Db      137 GMKLIEINTWLQVIPQLIARIDTHRQLVGQLIHQLLMDIGKNHPQALVYPLIVASKSASL 196

Qy     1246 ARHNAANKILLKNMCEHSNTLVQQAMMVSEBLIRVAILWHEMMWHEGLEEASRLYFGERRNVK 1305
                |  ::                           ::|:||
Db      197 ARNNAAFKILDSMRKHSPTLVEQAVMCSEBLIRVAILWHEQWHEGLEEASRLYFGDRNVK 256

Qy     1306 GMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYY 1365
            |||:|:||||| ||||||||||:|
Db      257 GMFEILEPLHAMLERGPQTLKETSFSQAYGRELTEAYEWSQRYKTSAVVMDLLDRAWDIYY 316

Qy     1366 HVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDENQPIIRIQSIAPSLQVITS 1425
            |||:|||||
Db      317 HVFQKISRQLPQLTSLELPYVSPKLMTCKDLELAVPGSYNPGQELRISIIKINLQVITS 376

Qy     1426 KQPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRY 1485
            |||||:||||
Db      377 KQPRKLCIRGSNGKDYMVLLKGHEDLRQDERVMQLFSLVNTLLLDDPTFRRNLAIQRY 436

Qy     1486 AVIPLSTNSGLIGWPHCDTLHALIRDYREKKILLNIEHRIMLRMAPDYDHLTLMQKVE 1545
            ||||||||||||||||||
Db      437 AVIPLSTNSGLIGWVPHCDTLHLTIRDYRDKDKKKVPLNQEHRTMLNFAPDYDHILTLMQKVE 496

Qy     1546 VFEHAVNNTAGDDLAKLIWLKSPSSEVWFPRRTNYTRSLAVMSMVGYILGLGDRHPSNLM 1605
            ||||
Db      497 VFEHALGQTQGDDIAKLIWLKSPSSEIWFERRNNYTRSLAVMSMVGYILGLGDRHPSNLM 556

Qy     1606 LDRLSGKILHIDFDGCFEVAMTREKPPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVME 1665
            ||| :|||||||||||||||||||||||||||||||:||:

Db      557 LDRMSGKILHIDFGDCFEVAMTREKPPEKIPFRLTRMLIKAMEVTGIEGTYRRTCESVML 616

Qy     1666 VLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYS--------AGQSV 1715
            ||| :||:|||||||||||||:|: |||

Db      617 VLRRNKDSLMAVLEAFVYDPLLNWRLLDVDKKGN------DAVAGAPGGRGGSGMQD 669

Qy     1716 EIILDGVELGEPAHKK---TGTTVPPSIHSPIGDGLVKPEALNKKAIQIINRVRDKLTGRD 1772
            ||| :::

Db      670 SLSNSVEDSLPMAKSKPYDPTLQGGLHNNVAD------ETNSKASQVIKRVXCKLTGTD 723

Qy     1773 FSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
            ||||: ||||||:||||:|||:|:::||||||||||

Db      724 FQTEKSVNEQSQVELLIQQATNNENLCQCYIGWCPFW 760


RESULT 12
Q8I6M7
ID   Q8I6M7;     PRELIMINARY;     PRT;   2604 AA.
AC   Q8I6M7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Phosphatidylinositol kinase-related kinase.
OS   Giardia lamblia (Giardia intestinalis).
OC   Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX   NCBI_TaxID=5741;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 50803;
RA   Morrison H.G., Zamora G., Campbell R.K., Sogin M.L.;
RT   "Inferring protein function from genomic sequence: Giardia lamblia
RT   encodes and expresses a phosphatidylinositol kinase-related kinase
RT   similar to yeast and mammalian TOR.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY095369; AAM33436.1; -.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Kinase.
SQ   SEQUENCE   2604 AA;  293344 MW;  7EFFB16F93BD59AF CRC64;

Query Match             21.9%;  Score 2064;  DB 5;  Length 2604;
Best Local Similarity   25.0%;  Pred. No. 1.2e-141;
  Matches  606;  Conservative 376;  Mismatches 750;  Indels 688;  Gaps  53;

Qy       27 LIRPYMEPILKALILKLKDPDFDPNPGVINNVLATIGELAQVSGLEMRKWVDELFIIIMD 86
            ::|  |     :      : :   |  ::  : :|  :   : ::
Db      236 IVLPYV-PMVSAVMGLISKDCDDEYLLITS-LTTLAELFTIGSTSMVHCLAEVIPLLIT 293

Qy       87 MLQDSBLLAKRQVALWTLGQLVASTGYVVEPYRKYPTLLEVLLNFLKTEQNGGTRREAIR 146
               :   |   :                   : ::|:|
Db      294 -CIKEETSKTLRLTTTRCFVRIVEATSFVVPWIYHDLFTIVFGILKNDTSQEIRLEAAK 353

Qy      147 VLGLLGALDP--YKH------------KVNIGMIDQSRDASAVSL--SESKS 182
            :: :||:|   ::
Db      354 LIITIGAIDPIFYQNGKPDLSALVPPTNLJCKCQSTDVFSLDYIMDESNIWVCHKGEGGA 413

Qy      183 SQDSSDYS-------TSEMLVNMGNLPLD------------EFYPAVSMVALMRIFRD 221
            :: :||:::|
Db      414 AQEASIKSHGKAVVSPITGHVVENYIMKLDSGSEVSANVSDFELTLNVIMSTLYSILSE 473

Qy      222 QSLS-IHHTMVQAITPIFKSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLV 280
            :: ||:|::| ::
Db      474 STLDMYHYDAVVVFRTIIQTECTAISLPMVPKIISILLENIQNCKVFMQEETMLRELILLI 533
```

```
Qy  1754 NKKAIQIINRVRDKLTGRDFSHDD------TLDVPTQVELLIKQATSHENLCQCYIGWCP 1807
              :: ||    ||                    ||  |||||||||:|||||  |||||||
Db  2635 STTDSIMETIKRKLDGTEFVHTDGSTPPEPLQVTEQLAMLTEQATSPLNLCQSYIGWCP 2693

Qy  1808 FW 1809
         ||
Db  2694 FW 2695
```

RESULT 10
O94189

```
ID   O94189      PRELIMINARY;   PRT;  1083 AA.
AC   O94189;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Phosphatidylinositol 3-kinase TOR1 (Fragment).
GN   TOR1.
OS   Cryptococcus neoformans var. neoformans.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC   Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX   NCBI_TaxID=40410;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B3501;
RX   MEDLINE=99262981; PubMed=10330150;
RA   Cruz M.C., Cavallo L.M., Gorlach J.M., Cox G., Perfect J.R.,
RA   Cardenas M.E., Heitman J.;
RT   "Rapamycin antifungal action is mediated via conserved complexes with
RT   FKBP12 and TOR kinase homologs in Cryptococcus neoformans.";
RL   Mol. Cell. Biol. 19:4101-4112(1999).
DR   EMBL; AF098973; AAD16274.1; -.
DR   HSSP; P42345; 1AUE.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IBA.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR000941; TPR-like.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Kinase.
FT   NON_TER       1      1
FT   NON_TER    1083   1083
SQ   SEQUENCE   1083 AA;  123973 MW;  0C48A17D758353A9 CRC64;

Query Match             28.7%;  Score 2703;  DB 3;  Length 1083;
Best Local Similarity   46.9%;  Pred. No. 2.6e-189;
Matches  547;  Conservative  189;  Mismatches  312;  Indels  118;  Gaps  16;

Qy  674 ESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDALVAY--DKKMDTN 730
            |:| |:||||| :||  | ||  | ||   |: |||||  ||| :||            
Db    6 EDLIDVNQKLQQSDAAWGTILEWAQTEMG-MTTEVEWYEKLGRWEEALQVWNERDADASTT 64

Qy  731 KDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAWGLGQWDSME 790
              |  :  |   |  :|||||||| :   | |          ::   |||: ||| |:|  
Db   65 FSEWEITEGKVTCLHAMGEWGQLSDFVQARWANRTAEEKKLLSPLARAAASWSLKQWDLMD 124

Qy  791 EYTCMIPRDTHDGAFYRAVLALHQDLPFSLAQQCCIDKARDLLDAELTAMAGESYSRAYGAM 850
           : : ::   : ||  :  |: :| :|  :  |  :| | ||  ||| | | :|: |||  |  
Db  125 DYISAMKGDGADRAFFKAIIAVHRNQIPAALKQISKARERLDPELTTLTGDSVGRAYDTV 184

Qy  851 VSCEMLSELEEVIQYK--LVPERREIIRQIWERLQGGCRIVEDWQKILMVRSLVVSPH 907
            :  ||:| |::  |||    :   |    || |||||:  || ||:|::|:  |      
Db  185 VRIQMLAEBLEIIAYKDHADEPARQEMQRQTWKKRLAGGCQRDVEWQRILQVRSIVLKPN 244

Qy  908 EDMRTWLKYASLCGKSGRLALAHKTLVLLLGVD-PSRQLDHPLPTVHPQVTYAYMKNMWK 966
          ||| ||:  | ||:|||| ||:|||   :|    :          | | ::  |||: |||   
Db  245 EDMDTWIEFADLCRTSDRLNLABKTLTSLVGFQVYPS--MEDTRGRAPPPIIFAYLRMAWA 302
```

```
Qy  967 SARKIDA----------FQHMQHFVQTMQQQA----------QHAIAIEDQQHKQELHKLMARCFL 10:2
           :||: |          |||                                              
Db  303 KNLQIDSREERYETLQHLRDFTDQLTDDVGIGARGPNGRLMLPDQKLYGSYTKLLAQCHV 362

Qy 1013 KLGEWQLNLQGINESTIPK-VLQYSAATEHDRSWYKAWHAWAVMNFEAVLHYHQNQAR 1071
         :||||| | :||  |  | |  ||||  |    || | || |  ||| | |:|       
Db  363 ELGQWQATLRESQGSADPSGILHDYSLATELDPEWYQAHTWALANFEVITQLEVSQQ-- 420

Qy 1072 DEKKKLRHASGANITNATTAATTAATTATATTTASTEGSNSESEAESTENSPTPSPLQKKVTE 1131
                   |                                                      
Db  421 --------GLSPIHFTT---------------------------------------- 429

Qy 1132 DLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQ 1191
                :|:|| ||||||| |||||  |||||||| |||||| ||| | | |||||| ||||  
Db  430 -------YIIPAVEBGFLKSISLSPGNSLQDTLRLLTLWFTYGYSSGVTAAVSQGLPTVN 481

Qy 1192 IDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAA 1251
         || ||:|||::|||| || : :  |: ::| | |:| ||||||| :||||| : || :|    
Db  482 IDVWLEVIPQIIARIQTPRQSIQQLIVQILHDIVKAHPQALIYPLIVASKSTVAARSTVA 541

Qy 1252 NKILKNMCEHSNTLVQQAMMVSEELIRVALLWHEMWHEGLEEASRLYFGERNVKGMFEVL 1311
           : : :| ||| :::|||| |||::|| ||: | | | | ||||:::||  ||   ::| | 
Db  542 QNITHKMREHSPKIVDQAELVSTELIRAAIIMHEWYDGLEEASKHYFGDHDIPGMLGVL 601

Qy 1312 EPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAMDLYHVFRRI 1371
         ||||  :| |||||:||||:|::| ||: || |   :   | ||||||| ||| | | | 
Db  602 EPLHRIVENGPQTLRETSFIQSFGHDLRIAREHLKRYRITQDGTEIQQAWDVYYSVFQRL 661

Qy 1372 SKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRK 1431
          |||  | :| ||||||||| | || | ||| | :| ||| | : :| |||| |||||: 
Db  662 GKQLKLLNVIELQYVSPKLMAVRDLDIAVPGTYQSGKPIIGIKNVIPTFKVIASKQKPRQ 721

Qy 1432 LTLMGSNGHBFVFLLKGHEDLRQDERVMQLFGLVNTLLANDFTSLRKNLSIQRYAVIPLS 1491
           : ||::|| : || ||||||||||||||||||||| |:||  | : :||||: :|:|| 
Db  722 CSMRGMDGKEYAYCLKGHEDLRQDERVMQLFGLVNTLLNNDHSAKRHLSIQRFSVTPLS 781

Qy 1492 TNSGLIGWVPHCDTLLHALIRDYREKKLLLNIEHRIMLRMAPD-YDHLTLMQKVEVFEHA 1550
            || ||| |  || : | || || | |  :|||::|  : : | :|| : :|||:|   | 
Db  782 PIAGLLGWVTHSDTIHVLIKQYRDQRKILVDIEHKLMQQMSDESYDSLPLLQKVEIFQYA 841

Qy 1551 VNNTAGDDIAKLLWLKSPSSEVWFDRRINYTRSLAVMSWVGYILGLGDRHPSNLMLDRLS 1610
          :| |:| : :|:||||| : :|:: | :| || :: | :||||||||||||| : |:: 
Db  842 LDNTTGQDLYRIIWLKSRNSDIWLERRTTYIRSLGLNSMVGYILGLGDRHPSNLLLDQIT 901

Qy 1611 GKILHIDFGDCFEVAMTREKFPPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREH 1670
         || :||||||||||| ::||:||:|||||||||: |||||| |:|:|| | |||||:|::  
Db  902 GKMVHIDFGDCFEVAQQRDKYPEKVPFRLTRMLIHAMEVCGITGNFSRQCEVSMEVLRDN 961

Qy 1671 KDSVMVLSAFVYDPLLNWRLMDTNTKGNKKSRTRTDSYSAGGSVEILDGVELGEPA--- 1727
         ::: |||:||||||| ||| |                 |  |||| | |:|||||:     
Db  962 RESLMAVLSAFVYDPLIAWRL----------------TATDKRPGGVG-EVKD---LDDPAVYG 1005

Qy 1728 ---HKKTGTTVPESIHSFTGDGLVKPEALNKKAIQINRVRKLTGRDPSHDDTLDVPT 1783
              ::: | :  |: :   :|    :::   |||:|| :||:|| | | ||: |||   
Db 1006 KQRKNKANETEIINDVEN------TEVKNDKGLQVIERVRRKLTGRDFKPDVVLDVKS 1057

Qy 1784 QVELLIKQATSHENLCQCYIGWCPFW 1809
         ||| :: ||:||||| | ::||| |||
Db 1058 QVEKLVVEARTKTENLCVAFLGWCSFW 1083
```

RESULT 11
Q960H0

```
ID   Q960H0
AC   Q960H0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SD02269p.
GN   TOR OR CG5092.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```

```
Q960H0          PRELIMINARY;       PRT;    760 AA.
```

```
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR009940; Prenyl_trans.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS00290; PI3_4_KINASE_3; 1.
KW   Hypothetical protein.
SQ   SEQUENCE  2695 AA;  306173 MW;  D15ADFA7B33825BA CRC64;

Query Match          36.4%;  Score 3425;  DB 5;  Length 2695;
Best Local Similarity  38.2%;  Pred. No. 9.3e-242;
Matches  780;  Conservative 337;  Mismatches 603;  Indels 322;  Gaps  44;
```

```
Qy     1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPLIKALILKLKDPDPDNPGVINNVLA    60
           ||| ||| |:|   ||| ||||::| :|| | |: ||| |:|  :  |  ||  ||:
Db   743 MQQSGQARLEQHSAKMIAQLAKQSPKFMRPYVGSIMIAMIPKLRNDQKYAE--VTAQVLN   800

Qy    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALMTLGQLVASTGVVVEYRK   120
           : |||| | |:  |||::||:| ||::|  ||| ||  |  ::| |: ||  |::||
Db   801 AVSEIAVIGGAEIVKNLKPLFEKLTHMINDSSLHKREAALRAIGGICRSTAYVVDPYRD   860

Qy   121 YPTLLEVLLNFLKTEQNGGTRREAIRVLGLLGALDPYHKHKVNIGMIDQSRDASAVSLSES  180
           ||:| |: |::| :|:| ||: |||||:|||  ||||||| |::| ::|  :|| : :||
Db   861 YPSLLDDLLRIIKTVMSNTMRREAIKTLGIIGAIDPYTHKVFTGSV-QSSTAISTALSLP  919

Qy   181 KSSQDSSDYSTSEML--VNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFI   238
             |: | ||:|:||    |:|| |  ||| || || |: ||: ||:|:  |||||:|
Db   920 ISETDSKD-PRQDIIHWPNYEKCTLEEFYPAITIANLMLMWQDED-SQSYAEIAQAIVTI   977

Qy   239 FKSLGLKCVQFLPQVMPTFLNVIRVCDGA-------IREFLFQQLGMLVSFVKSHIRPY   290
           |:|||  |||| |::|| | ::|  |      ||| |:||:| ||| |:|
Db   978 FRSLGDMAPLYTEQVIP---RLIEVCRRAITSSNRANLREFFLQQLANPVAIIRKHAAPY  1034

Qy   291 MDEIVTLMREFWVMNTSIQSTILLIIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGR  350
           | :| | :   :|| :  :||::|:|: |:|  ||  | ||  ::| : |:   ||
Db  1035 MPAIFTIIADAWKEDISVXMVVIEVLTDMGTAIGNDFSKYTGELIIPYLLTVLQTDKTKER  1094

Qy   351 IVSIKLLAAIQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAARLETVDRLTTESLDFTD   410
           |:|::|  : |:|| |||  : |: | :|| |    | ||: ::  ::| |::|||::||
Db  1095 VLTVKVMRESIQKLTHCIVQELHLVLPPLLIILDDFSLKLSIRNTALSTVLHMTQQVDVSA   1154

Qy   411 YASRII--HPIVRILDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVL----VRH   462
           || :|   |||||:|||:    |::|| |::||| ||: : |:|||| |   |:|
Db  1155 YAPRMMQQSWHHNI-----STAEMRDKLLLIEIKQLGKFPDIFKRGVDQKLRDYNLDK   1209

Qy   463 RINHQRY---------DVLICRIVKG-------------YTLADEEEDPL    490
           | :| | ::   |    :|::|:|   ||  ||:| |
Db  1210 SVHYEQYRKLAQRAQMSRDVLTSSVFAGSNGNIQYSSTQAGMRGQANNYANND--------  1263

Qy   491 IYQHRMLRSGGQDALAS---------------------GPVETGPMKKLHVSTINLQ   526
           | |:  | || |:||              ||:| || | |:|
Db  1264 --LHERLMNGSIDSGASRQDNRDDYRYGVEEKISEEVPKVAPTTARPTSELVTVQITKQ  1321

Qy   527 K-------AWGAARRVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLF   578
           |       ||:| |:: |    || :|:|:  |:|::|  |:  || || |::|||| |
Db  1322 RLNKDALMPQWKNENLTSKDEWLIQWLMKIRIGFLTYGSSPSLRAASSLGDQHPHLARDLF  1381

Qy   579 NAAFVSCWSELNEDCQQDELIRSIELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDN   638
           ||| | || ::|| ::| | |::|| | :|::|  || |||| |||:|:  | ||:  ::|
Db  1382 PAAFMSVWTELDSDVCNDL-TSCLLRAISTGIPELIQTLNLAEFMDHSEKGPLPISHD-  1439

Qy   639 GIVLLGERAAKCRAYAKALHYKELEFQKGP----------------TPAILEBSLISINNKLQQ   685
           ||:|||| || || | ||||| |||  ||      ||| ||:| |:|:||
Db  1440 ---VLGRWAEQTKAFAKACRYKEMSVLKKSSSMQTTFTRKVKLEPNDCQSLITYANKLNV  1496

Qy   686 PEAAAGVLEYAMKHFGELEIQAIWYEKLHEWEDALVAYDKKMDTNKDDPELM----737
           | ||||| || |  |  |:||:|| |:||| ||| | || ||| ||| |:|
Db  1497 QEPAAGVVRYAERNEMNFQMRGRWYEKLNEWEKALGAYELEKEKKSSCPNLQVYDEKDHL  1556

Qy   738 --------------LGRMRCLEALGEWGQLHQQCCEKWT------LVNDETQA------K   771
                        || ||| || ||:| | ||: : |   : |:
Db  1557 MTPEEAATAEEARMEHMRCLEALGRWDELNSKSV-VWADQRGNRNDSVRDEINKKQLDHK  1615
```

```
Qy   772 MARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLL   831
           | | |||| |:|  |::|:||  :|  |  |   ||: | |   :::|:|
Db  1616 MAVIAARGAWAVDNWERMADYVSVISENTQDGAMLRAVVAVANDENTKAMGLIEKVREMI  1675

Qy   832 DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIRQIWERLQGCQRIVE   891
           |:||:|| |: |: ||   :: |:||: ::|||:| ::|:|| |:|:   :|:|:::|
Db  1676 DSELTAMANESYERAYIPMWSVQQMAELEEAIEYKTRPRRPRIALLWSRLQGCRRNVE  1735

Qy   892 DWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDH-PLP   950
           |:| |||| |:|:|| ||: ||||: | :|| : || |:|: :||| :||| | |:| |||
Db  1736 QWQRLIMLRGLVLSPQEMPLRVKFSSMCRKQGKNSMSRAVLRELSLPANSDLVRAKAP  1795

Qy   951 TVHPQVTYAYMKNMWKSARKIDAFQEMQHFVQTMQQQAQHAIATEDQQ---HKQELHKL  1006
           | || : |: |:||:|:|::|:| ::| || || :: ::|| ||:|:|
Db  1796 FDKPLLVLALAKQLYQDDHKDEAIRALEDILANEHWNKRINPIPKATGRELIPPSTKEPARI  1855

Qy  1007 MARCFLKLGEW---------QLNLQ--------GINESTIPK------VLQIYSA  1038
           ||: ||:|| :      ||:|         |:|: ||: :|:|||||:||
Db  1856 CAKVLLKIGEWTELKSKTSNNMVGELSFVRQQVSPQVRTKESRTPETIAFENTINYQQ  1915

Qy  1039 ATEHDRSWYKAWHAWAVMNFEAVLHYK-HQN---QARDEKKKLRHASGANIITNATTA  1091
           |:|  |:||:| :|: |:| | || |:     ||:|| |::| ||: |:::|::||
Db  1916 ATQYDPGWHKWVHKLASTHFYAVCRERPHPTTVISPPQQPQQPKKMHIPP---VTRATSP  1972

Qy  1092 ATTAATATTTASTEGSNSESEABSTENSPTPSPLQKKVTEDLSKTLLMVTVP-------  1143
           |::|    |: :  : ::|  |:|: |::| :|:|| |:|||::|::|::|
Db  1973 PPPA------------QKSPQPAPFH-SITEPLS-VIIDYPVPPLGSLVG           2009

Qy  1144 --------------------------------AVQGPFRRSLSLRGNN            1159
                                           || | ::| :| ::|
Db  2010 LPPMPAYLSSNSSLPPQHHVSPLSNDSPSNSAENKLYLKHAAHAVRCPAKALMCSPGSR  2069

Qy  1160 LQDTLRVLTLWFDYG--HWPDVNEALVEGVKAICIDTWLQCVIPQLLARIDTPRPLVG-RL  1216
           |:|||| ||:|||  ||| || |||  || : || |:|   |  :| :|  |:|   | |
Db  2070 LEDTLRLMQLWFDHGDDKDQDVYFALTETIFDLPVTTWLEAIPQLMARIDCPDDQKSVQL  2129

Qy  1217 IHQLITDIGRYHPQALIYPLTVASKSTTTARHNAANKILKMCEHSNTLVQQAMMVSEEL  1276
           :  |:|| |:|  || ::|| ||:| || :| || |:||:| |: | | ||: |:|:||
Db  2130 VLRVLCEIARHPQAVIVALTVASRSKDVHRSKNAGTVLEKMMEYHSKLVREASLVTEEL  2189

Qy  1277 IRVAILWHEMWHEGLEEASRLYPGERNVKGMFEVLEPLHAMMERG-PQTLKETSFNQAYG  1335
           :| || |:||:|| || | |:|:| ||::  || | |:| ::| : |  ||:|| :
Db  2190 VRCAILWHEQWHEDALDDASRVYFHDNNVQAMFDALRNMNDLMQKGAPTTMKEASFQQTYS  2249

Qy  1336 RDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTSLELQYVSPKLLMCRD  1395
           |||:||:|||| || || |:|:|:  ||: ||  |  :|  :|  |:|  |||: |:|
Db  2250 SDLKEAGRYVQAFESSGNVKDLNQAWEIYCSVPFKKLRDQLATLNSLDLVYYSPNLVSAKD  2309

Qy  1396 LELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQD  1455
           ||:|||| ||||:|  |:|: ||:|| |:|:||| | || ||||  |: || |||| ||
Db  2310 LELVVPGTYDPSAPIVSI-QSFSSKMNVITSKQRPRKMVIRGSNGLDYQFLLKGHEDPRQD  2369

Qy  1456 ERVMQLFGLVNTLLANDPTSLRKNLSIQRYAV:PLSTNSGLIGWVFHCDTLHALIRDYRE  1515
           ||||||||||||||| :||:|||||| |||| |:|:::||| |||||||:|:|:||||:|
Db  2370 ERVMQLFGLVNTLLANNSETCRRNLITIQRYSIVALSKDSGLIGWVPNCDTLHTLVKSYRE  2429

Qy  1516 KK-KILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWF  1574
           || | |::: |::| :|||| || | |:|:| | |:: :|:||| :|::||:|||||||||
Db  2430 KKAKIPLSIEHKTIQKLSLETEHLIIMQKLQLFESALSVTQGEDLRHVLWLKSPSSEVWF  2489

Qy  1575 DRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFPGDCFEVAMTREKFPEK  1634
           ||||||||| ||| ||||||||||||||||| ||:|| : |||| || || | || ||:
Db  2490 DRRTNYTRSVACMSMVGYILGLGDRHPSNLMLDRLTGKVVHIDFGDCFEVAMLREKFPER  2549

Qy  1635 IPFRLIRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLNWRLMDT  1694
           |||: |: ||||||||| |||| | | :| |||::| ::| || |||  |::| ||:
Db  2550 VPFRLTRMLINAMEVTGLDGVYNYTAERVLKMLRTNQESLLAVLEAFVYDPVINWRLVE-  2608

Qy  1695 NTKGNKRS-RTRTDSYSAGQSVEILLDGVELGEPAHKKTGTTVPESIHSPIGDGLVKPEAL  1753
           |  |:|:  : |:  :|| :
Db  2609 --GMKKDPKIRKD--TGGR------------------------QNMAGAVLPSS------  2634
```

```
Db  1155 YAPRMMQSWHHNI-----STABMRDKLLLLLIEIIKQLGKFPDIFKRGVDQKLRDYNLDK 1209
Qy   463 RINEQRY-------DVLICRIVKG--------YTLADEBEDPL                     490
Db  1210 SVHYEQYRKLAQRAQMSRDVLTSSVFAGSNGNIQYSSTQAGMRGQANNVYANND----    1263
Qy   491 IYQHRMLRSGQDALAS-----------GPVETGMKKLHVSTINLQ                  526
Db  1264 -LHERLMNGSIDSGASRQDNRDDYYRYGVEEKISEVPKVAPTTARPTSELVTVQITKQ    1321
Qy   527 K-------AWGAARRVSKDDWLEWLRRLSLELLKDSSPSLRSCWALAQAVNPMARDLF     578
Db  1322 RINKDALMPQWKNENLTSKDEWLQWLMKIRIGPLTYGSSPSLRAASSLGDQHPHLARDLF  1381
Qy   579 NAAFVSCWSELNEDQQDELIRRSIELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDN 638
Db  1382 PAAFMSVWTELDSDVQNDL-TSCLLRAISTGIPELIQTIINLARFMDHSEKGPLPISHD-  1439
Qy   639 GIVLLGERAAKCRAYAKALHYKELEFQKGP----------TPAILESLISINNKLQQ     685
Db  1440 ---VLGRWAEQTKAFAKACRYKEMSVLKKSGSMQTTFTRKVKLEPNDCQSLITYANKLNV  1496
Qy   686 PEAAAGVLEYAMCHFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELM--------  737
Db  1497 QEEAAGVVRYAERNEMNFQMRGRWYEKLNEWEKALGAYELEEKKKSSCPNLQVYDEKDHL  1556
Qy   738 -----------LGRMRCLEALGEWGQLHQQCCEKWT-------LVNDETQA-----K     771
Db  1557 MTPEEAATAREARMHEMRCLEALGRWDELNSKSV-VWADQRGNRNDSVRDEINKKQLDHK  1615
Qy   772 MARMAAAAAWGLGWDSMEEYTCMIPRDTHDGAFYRAVILAHQDLFSLAQQCIDKARDLL   831
Db  1616 MAVIAARGANAVDNWERMADYVSVISENTQDGAMLRAVVAVHNDENTKAMGLIEKVREMI  1675
Qy   832 DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVE  891
Db  1676 DSELTAMANESYERAYIPMVSVQQMAEBLEEAIEYKIRPERRPRIALLWSRRLQGCRRNVE 1735
Qy   892 DWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDH-PLP   950
Db  1736 QWQRLIMLRGLVLSPQEMHPLRVKFSSMCRKQGKNSMSRAVLRELLSLPANSDLVRAKAP  1795
Qy   951 TVHPQVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQQAQHAITATEDQQ---HKQELKHL  1006
Db  1796 FDKPLIVLALLAKQLYQDDHKDEAIRALIEDLANEHWNKRINPIPKATGREELIPPSTKEPARI 1855
Qy  1007 MARCFLKLGEW-------QLNLQ------GINSSTIPK------VLQIYSA          1038
Db  1856 CAKVLLKLGEWTELKSKTSNMQVGELSFVRQQVSPQYRTKESRTPETIAPENTINYYQQ   1915
Qy  1039 ATEHDRSWYKAWFAWAVMNPFEAVLHYK-HQNQARDEKKKLRHASGANITNATTAATTAAT 1097
Db  1916 ATQYDPGWHKVWHKLASTHFYAVCRERPHPTTVISPPQPQQPEKKMHIPPVTRA--TSPP  1973
Qy  1098 ATTTASTEGSNSESEAASTENSPTTSPL----------QKKVTEDLS--            1134
Db  1974 PPAQKSPQPAPFHSITEPLSDYPVPPLGSLVGLPPMPAYLSSNSSLPPQHHHVSPLSND   2033
Qy  1135 ------KTLLMYTVPAVQGFFRSI4LSRGNNLQDTLRVLTLWFDYG--HWPDVNEALV    1184
Db  2034 SPSNSAENKLYLKEAHAVRCPAKAIMCSPGSRLEDTLRIMQLWFDHGDDKDQDVYFALT   2093
Qy  1185 EGVKAIQIDTWLQVIPQLIARIDTPRPLVG-RLIHQLLTDIGRYHPQALIYPLIVASKST  1243
Db  2094 ETIFDLPVTTWLEAIPQLMARLLDCPDDQKSVQLVLRVLCEIARHRPQAVIYALTVABRSK 2153
Qy  1244 TTARHNAANKILKNMCEHSNTLVQQAMWVSEELIRVALIWHEMWHEGLEEASRLYFGERN  1303
Db  2154 DVHRSKNAGTVLEKMMEYHSKLVREEASLVTEBLVRCAILWHEQWHDALDLDDASRVFHDNN 2213
Qy  1304 VKGMPEVLEPLHAMMERG-PQTLKETSPNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWD  1362


Db  2214 VQAMPDALRNMDLMQKGAPTTMEHSFQQYSSDLKEAGRYVQAPESSGNVKDLNQAWE     2273
Qy  1363 LYYHVRRISKQLPQLTSLELQYVSPKLLMCRDLEBLAVPGTYDPNQPIIRIQSIAPSLQV  1422
Db  2274 IYCSVFKKLRDQLATLNSLDLVVYSPNLVSAKDLELVVPGTYDPSAPIVSIQSFSSKMNV  2333
Qy  1423 ITSKQRPRKLTMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSI   1482
Db  2334 ITSKQRPRKMVIRGSNGLDYQFLLKGHEDPRQDERVMQLFGLVNTLLANNSETCRRNLTI  2393
Qy  1483 QRYAVIPLSTNSGLIGWPHCDTLHALIRDYREKK-KILLNIEHRIMLRMAPDYDHLTLM   1541
Db  2394 QRYSIVALSKDSGLIGWPNCDTLHTLVKEYREKKAKIPLSIEHKTLQKLSLETEHLTLM   2453
Qy  1542 QKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHP  1601
Db  2454 QKLQLFESALSVTQGEDLRHVLWLKSPSSEVWFDRRTNYTRSVACMSMVGYILGLGDRHP  2513
Qy  1602 SNLMLDRLSGKIIHIDFDGCFEVAMTRKFPEKIPPRLTRMLTNAMEVTGLDGNYRITCH   1661
Db  2514 SNLMLDRLTGKVVHIDFDGCFEVAMLREKFPERVPPRLTRMLINAMEVTGLDGVYNYTAE  2573
Qy  1662 TVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRS-RTRTDSYSAGQSVEILDG  1720
Db  2574 RVLKMLRTNQESLLAVLEAFVYDPVINWRLVE---GMKKDPKTRKD--TGGR          2620
Qy  1721 VELGBPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDD-   1777
Db  2621 -----QNMAGAVLPSS-----------------STTDSIMETIKRLDGTEFVHTDGST  2657
Qy  1778 ---TLDVPTQVELLIKQATSHENLCQCYIGWCPFW                          1809
Db  2658 PPEPLQVTEQLAMLTEQATSPLNLCQSYIGWCPFW                          2692


RESULT 9
001438
ID   001438      PRELIMINARY;      PRT;   2695 AA.
AC   001438;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein B0261.2a.
GN   B0261.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Scheet P. Gattung S.;
RT   "The sequence of C. elegans cosmid B0261.";
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U97016; AAN4885.1; -.
DR   WormPep; B0261.2a; CE32559.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
```

```
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Scheet P., Gattung S.;
RT   "The sequence of C. elegans cosmid B0261.";
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U97016; AAN84886.1; -.
DR   WormPep; B0261.2b; CE32560.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008940; Prenyl_trans.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Hypothetical protein.
SQ   SEQUENCE  2692 AA;  305847 MW;  80B31C5469D59261 CRC64;

Query Match             36.4%;  Score 3427.5;  DB 5;  Length 2692;
Best Local Similarity   38.6%;  Pred. No. 6.1e-242;
Matches  778;  Conservative 337;  Mismatches  629;  Indels  271;  Gaps  41;

Qy         1 LEHSGIGRIKEQSARMLGHIVSNAPRLIPYMEPILKALILKLKDPDPDNPGVINNVLA   60
             :|||||| ||  | :::|||:   | : :|| ::  :  | ::||: |   ::  |
Db       743 MQQSGQARLEQSAKMIAQLAKQSPKFMRPYVGSLMIAMIPKLRNDQKYAE--VTAQVLN  800

Qy        61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVVEPYRK  120
             |:||| |||||   :|:  | :|||::||  ::| |:::|::| |: |::||:: |||:|
Db       801 AVSEIAVIGGAEIVKNLKPLFKLTHMINDSSSLHKREAALRAIGGICRSTAVVVDPYRD  860

Qy       121 YPTTLBVLLNFLKTEQNGSTRREAIRVLGLLGALDPYHKVKYNIGMIDQSRDASAVSLSES  180
             | :: |  :||  :   ||  |||||||:|||:||||||||| ||||| |   :   ::|
Db       861 YPSLLDDLRIIKTVMSNTMRRRAIKTLGLIGAIDPYTHKVFTGSV-QSSTAISTAISLP   919

Qy       181 KSSQDSDYSTSEML--VNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFI  238
             | |:  :| :   :     | : | :: | |  :|  | |:     : ::  : ::|
Db       920 ISETDSKD-PRQDIIHWFNYEKCTLEEFYPAITIANLMLMMQDED-SQSYAEIAQAIVTI  977

Qy       239 FKSLGLKCVQFLPQVMPPTFLNVIRVCDGA----------IREFLFQQLGMLVSFVKSHIRPY  290
             |:|||||||  :|||||||||| ||||:||            |:|::| |:::| :||:|||
Db       978 FRSLGDMAPLYTEQVIP--RLIEVCRRATESSNRANLREFLLQQLANFVAIIRKHAAPV  1034

Qy       291 MDEIVTLMREFWVMNTSIQSTILLIEBQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGR  350
             | :: :|:|   |::    : : : || |::|  |  :|  | ||:  | |:| :|| |
Db      1035 MPAIFTIIADAWKEDISVKMVVIEVLTDMGTAIGNDFSKYTGELIPYLLITVLQTDKIKER  1094

Qy       351 IVSIKLLAAIQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTD  410
             :||:||| |: :||::: :  :|| ||| |::|::| :  | :: :| |: :: :|| :
Db      1095 VLTVKVMESIQKLTHCIVQHLHIVLPPLLILLDDFSLKLSIRNTALSTVLHMTQQVDVSA  1154

Qy       411 YASRII----HPIVRTLDQSPELRSTAMDTLSSIVFQLGKKYQIFIPMVNKVL----VRH  462
```

```
Query Match          40.9%;  Score 3851;  DB 10;  Length 2513;
Best Local Similarity 42.5%;  Pred. No. 3.7e-273;
  Matches  830;  Conservative  332;  Mismatches  515;  Indels  276;  Gaps  38;

Qy     4 SGIGRIKEQSARMLGHLVSNAPRLIRPYMEPTIKALLILKLKD-PDPDNPGVINNVLATI   62
           :.|::.|.|| |.:|| ::|  :|:.|  :||.|.|:||.|:||| :|:::.:.|:::|
Db   677 SADNKCRESSAKLLGCLVRNCERLILPVVAPVQKALVARLSEGTGVNANNNIVTGVLVTV  736

Qy    63 GELAQVSGLEMRKWVDELFIIIMDLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRKYP  122
           :|:|.|:|:|.||.:::..::.:|:||:.:||:||.:|||.:|::.|.|:|.:|::||.
Db   737 GDLARVGGLAMRQYIPELMPLIVEALMDGAAAVAKREVAVSTLGQVVQSTGVVVTPYKEYP  796

Qy   123 TLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYHCKVNIGMLDQSRDASAVSLSEKS  182
           ||.||:|.|| |:.|.           ::|||:|:|::|:||:||  |  |.:|:|:|
Db   797 LLIGLLLKLJKGDLVWSTRREVLKVLGIMGALDPHVIK-------RNQQSLSGSHGEV  847

Qy   183 SQDSSDYSTSEMLVNMGNLPL------DEFYPAVSMVALMFPRDQSLSHHHTMVVQ   233
           ||.||.||:.|::|.|.||         :|:.|:.|:|| :|.|:||::.::.|.:||:
Db   848 PRGTGD--SGQPIPSIDELPVELRPSFATSEDYISTVANSLMRILRDASLLSYHKRVVR  905

Qy   234 AITFIFKSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDE  293
           :: .|::  :|.::||.:.|       :|:.:.:..|.|.:.:.||||.|||:|::|:
Db   906 SLMIIFK----------------VLPELFHTVRTSDENLKDFITWGLGTLVSIVRQHIRKYLPE  953

Qy   294 IVTLMREFWVMNT----------------SIQSTIILLIBQIVVALGGEFKLYLP   332
           :| |::||.|.|                      .||::||.|.:|::|||.||.
Db   954 LLSLVSELNSSFTLPGPIRPSRGLPISYGEYFASLTDNQVLHLLEHLCLALNDEPRTYLP 1013

Qy   333 QLIPHMLRV--------------FMHDNSPGRIVSIKLLAATQLFGANLDDYLHLLLPPIVKL  381
           :|     .:|                 :.||  :.::|.||::.|||||:.|:|.||||:||:|
Db  1014 VILPCFIQVLGDAERFNDYTYVPD---------LHTLEVFGGTLDEHMILLPALIRL  1063

Qy   382 FDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIVRTLD-QSPELRSTAMDTLSSI  440
           :|.|:.::| .: ::|:.|.:|::|||||.|  :|||:||.|  |||| |: ||.||.|.|
Db  1064 FKV-DAPVAIRRDAIKTLTRVIPCVQVTGHISALVHHLKVLVLDGKNDELRKDAVDALCCL 1122

Qy   441 VFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICR------IVKGYTLADEE-------  486
           :.|:|::||:|:|:|:|:|..|:|:|:|:|| |           :|||.|||
Db  1123 AHALGEDFTIFIBSIHKILLKHRLRHKFEEIHARWRRREPLIVA--TTATQQLSRRLPV 1180

Qy   487 --EDPLIYGHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDWLEW  543
             :||||.|:|:|.|:|.|.||.||.|.||  ||||:|:|.|::|.||.||:|:||:||
Db  1181 EVIRDPVIENE------------IDPFEEGTDRNHQVNDGRLRTAGEASQRSTKEDWEEW 1228

Qy   544 LRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRSIEL  603
           |.|||::|||||||||||| | |::|.|::| |:||.||.|||::|:.| .|::|:||:|
Db  1229 MRHFSIELLKESPSPALRTCAKLAQLQOPFVGRELFAAGFVSCWAQLNESSOKQLVRSLEM 1288

Qy   604 ALTSQDI-AEVTQTLLNLAEFMEHSDKGPLPLRDNGIVLLGERAAKCRAYAKALHYEKL  662
           |:|.::| ::|.:||.|||||  :||.|.::|::.||::|||:.||.||||:||.||:.|
Db  1289 AFSSPNIPPELLATLLNLAEFMEHDEK-PLPI---DIRLLGALAEKCRVFAKALHYKEM 1343

Qy   663 EFQKGPTP--------AILESLISNNKLQOPEAAAGVLEYAMKHPGELBIQATWYEKLH  714
           :|:||.|          ::|.::|:.:|:|::|:|.|:||:||:|:|||.||.|.|::||
Db  1344 EFE-GPRSKRMDANPVAVVEALIHINNQLHQEAAVGIITYAQQHL-DVQLKESWYEKLQ  1401

Qy   715 EWEDALVAYDKKMDTNKDDPELM----LGRMRCLEALGEWGQLHQQCCEKWTLVNDETQA  770
           ||:||:|:|:|.|::|.|||:||     |.||.||:|||.|||:|.|| |:::.|.|::|
Db  1402 RWDDALKAYTLKA-SQTTNPELVLEATLGQMRCLAALARWEELNNLCKEYWSPAEPSARL 1460

Qy   771 KVARMAAAAAWGLGQWDSMEEYTCMI----------------PRDTHDGA----FYRAVLALH  813
           |.||:|:||.:|| |||:|:|:.|:.|                  .|.||::    :|:||.|:
Db  1461 EMAPMAAQAAAWNMGEWDQMAEYVSRLDDGDETKLRGLASPVSSGDGSSNGTFFRAVLLVR 1520

Qy   814 QDLFSLAQCCIDKARDLLDAELTAMAGESYRAYGAMVSCHM.SRLEEVIQYKLVP------  869
           :|||:|||| :||:.| | :.|::.|:||:|||::|||| | |:||||:||:||
Db  1521 RAKYDEARBYVERARKCLATELAALVLESYERAYSNMVRVQQLSELEEVIEYYTLPVGNT 1580

Qy   870 ---ERREIIRQIWERLQGCORIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRL  926
              |||::|:|:||.|.|.|| ||||.|::|:|.:|||.|||:||||||||.||..||:.|
Db  1581 IAEERALLRNMWTQRIQGSKRNEVWQALLAVRAIVLLPPTEDVETWLKFASLCRKSGRI 1640

Qy   927 ALAHKTLVLLLGVDPDSRQLDHPLPTVHPQVITYAYMKMW---KSARKIDAFQEMQHFVQT  983
```

```
RESULT 7
Q9LPM4
ID   Q9LPM4          PRELIMINARY;     PRT;  2513 AA.
AC   Q9LPM4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   F2J10.9 protein.
GN   F2J10.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Sakano H., Liu S.X., Yu G., Lee J., Lenz C., Pham P., Toriumi M.,
RA   Chin C., Chiou J., Choi E., Chung M., Gonzalez A., Howng B., Liu A.,
RA   Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA   Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA   Miranda M., Nguyen M., Palm C., Shinn P., Southwick A., Davis R.W.,
RA   Ecker J.R., Federspiel N.A., Theologis A.;
RT   "The sequence of BAC F2J10 from Arabidopsis thaliana chromosome 1.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC015445; AAF76442.1; -.
DR   PIR; G96536; G96536.
DR   HSSP; P42345; 1FAP.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS00290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE  2513 AA;  282911 MW;  A4B9740321AC5261 CRC64;
```

```
Qy  1757 AIQIINRVRDKLTGRDFS---------------HDDT------------------------LDVPTQVELLIKQATS 1794
         |:.::|..: |||||||:                :|:                        |||||||||||||||||
Db  2407 AVVWARMSHKLTGRDFSSSAIPSNPIADHNNLLGGDSHEVEHGLSVKVQVQKLINQATS 2466

Qy  1795 HENLCQCYIGWCPFW 1809
         |||||| :|:|||||
Db  2467 HENLCQNYVGWCPFW 2481


RESULT 6
094188
ID   094188;              PRELIMINARY;       PRT;   2360 AA.
AC   094188;
DT   01-MAY-1999 (TrEMBLrel 10, Created)
DT   01-MAY-1999 (TrEMBLrel 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel 24, Last annotation update)
DE   Phosphatidylinositol 3-kinase TOR1.
GN   TOR1.
OS   Cryptococcus neoformans var. neoformans.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC   Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX   NCBI_TaxID=40410;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H99;
RX   MEDLINE=99262981; PubMed=10330150;
RA   Cruz M.C., Cavallo L.M., Gorlach J.M., Cox G., Perfect J.R.,
RA   Cardenas M.E., Heitman J.;
RT   "Rapamycin antifungal action is mediated via conserved complexes with
RT   FKBP12 and TOR kinase homologs in Cryptococcus neoformans.";
RL   Mol. Cell. Biol. 19:4101-4112(1999).
DR   EMBL; AF098972; AAD16273.1; -.
DR   HSSP; P42345; 1AUE.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Kinase.
SQ   SEQUENCE   2360 AA;  267216 MW;  ED4A1059B1AA2B2A CRC64;

Query Match             40.9%;  Score 3852;  DB 3;  Length 2360;
Best Local Similarity   43.6%;  Pred. No. 2.8e-273;
Matches  811;  Conservative 320;  Mismatches 553;  Indels 174;  Gaps  29;

Qy     1 LEHSSIGIRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKPDPDDPNPGVINNVLA 60
         |||  |  |:||: || ||:||:||||   | |||| : |:|| | | |    |||||||
Db   628 LEYSTNARQKEESAKLLCLMIGAAAGLVKSYAPTILSVLLRTAS--SPESSIGVQAECLK 685

Qy    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
         |||  |  |:|  |: ::|:|::  ||| :||||   |:||::|:|| ::|:: |  :|
Db   686 CIGELARVAGEELVPSVRAILDLVIEMLNDQSSPAKRDTALKTLGQIASNTGEVIKPYTD 745

Qy   121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQ------SRDAS 173
         |||| || | ::|::|: |||| |||:|||||| | |: ||||::|||      ||| |
Db   746 YPQLMGVLFKFLRMEANSSVRQEIYKTIGMGALDPFKHKTLLGDVDDPIDEGTSRVND 805

Qy   174 AVSLSESKSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSWVALMRIFRDQSLSHHHTMVVQ 233
         | | | :  :|::|::  |:| ||:  | ||:     | :     |: ||::||:|: :|
Db   806 IVLNQHNSS---------VN------DEFPQTVVIHSLVNVLHDSTYKDHY-QAVE 846

Qy   234 AITFIFKSIGLKCVQFLPQVMPTFLNVIRVCDGAIREFLPQQLGMLVSFVKSHIRPYMDE 293
         | ||:|:|| | |  ||| : ||:||: |::| |: :||  |:|::|:|:|::|||:|:
Db   847 AIMMIFRTQRLRCVNFLPQIVPAFLNVIRIAHSSRTELYLKQLAQPITIVKLHIRNVLND 906

Qy   294 IVTLMREFWVNNTSQSTIILLIEQIVVALGGEFKLYLPLIPHMLRVFMHDNSPGRIVS 353
         || ||| : ::|::|:|  || ::|:   || |||   | |:  | |:|:: || :
Db   907 VFDLIHEFWPNSTLQITIISLVEAIAKAVEGEFKAYLPKLLQQILRSFPGDLSAKHLPE 966
```

```
Qy   354 IKL------LAAIQLFGANLDDYIHLLLPPIVKLFDAPBAPLPSRKAALEVDRLTESLD 407
         :||      |||:|||:  |::: | || |::|:  :  | :|  : ||:|::|:::
Db   967 LKLNTLLQLIKAFYVFGESIEDYLHLVLPVIVRSFENPAAPDSLRIAALRTTGQLCRKVN 1026

Qy   408 FTDYASRIIHPIVRTL-DQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVANKVLVRHRINH 466
         ||:||:||||:||:  |:|: ||:  | :|:  | || |:||  : | |:||:||:|:|:|
Db  1027 FSDHASQIIHPLVRTLGNSSEELRQTAMETLCVLVLQFGPDYAIFIPMVNKALVENKISH 1086

Qy   467 QRYDVLICRIVKGYTLADEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQ 526
         |||| ||   ||:| ||:|:|: ||| | ||:|: | ||:|:| ||||  |||:|||:||
Db  1087 PGYRALITKLLNRERLPPD-----LGPVERYASDSAAEASAPPPV------ALKNQQALK 1136

Qy   527 KAWGAARRV--SKDDWLEWIRRLSLELLKDSSPSLRSCWALAQAVNPMARDLFNAAFV 583
         ||:|||:|   |||  ||||  | ||: |||::||  |:| |||  ||:|::|| |:|
Db  1137 LAWDCSHLLNTSSRTEWISWIIGLGHEMMRRSPSQAIRAARSLALSVAFTKELFNVAFY 1196

Qy   584 SCWSELNEDQQDELIRSIELALTSQDI-AEVTQTLLNLAEFMEHSDKG-PLPLRDDNGIV 641
         |||:|||::|| |::|:|:|:|:|:|  | |:|||:|:: |::| | | ||| |: :::|
Db  1197 SCWQELFSYQEDIWHNLDRAIKKDDVPGDVVNMILGATVQFLEHDEKEVAIESR------ 1250

Qy   642 LLGERAAKCRAYAKALHYKELEFQKGPTPAILESLSINNKLQQPEAAAGVLEYAMKHFG 701
         :|||  |  :|: || |:|::: |: |:  :|||| | ||::||:|||||:||||:|:|
Db  1251 VLGSVAANYQALAVALVHYKEQEFFLDPSKEVIEDLIDVNQKLQQSDAAWSTLEWAQTEMG 1310

Qy   702 ELEIQATWYEKLHEWEDALVAY--DKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCC 758
         : ::|: ||:|| || | |:    |: |:   | ||: : | ::  :|| |  |:::|
Db  1311 -MTTEVEWYEKLGRWEEALQVWNERDADASTTFSEWEITBGKVTCLHAMGEWEQLSDFVQ 1369

Qy   759 EKWTLVNDETQAKVARMAAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVLAJHQDLFS 818
         :|| |:|:||: | | ||  :|:|  |  ||: ||:|  |   |   :|::| : |::  :
Db  1370 ARWANRTAEEKLLSPLAAAASWSLKQWDLMDDYISAMKGDGADRAFFKAILAVHRNQIP 1429

Qy   819 LAQQCIDKARDLLDALTAMAGESYSRAYGAMVSCHMLSELEVIQYK---LVPERREII 875
         ::::||:|| || | : |||: :| : || ::: :: |::|:||     |:|: |:||
Db  1430 AALKQISKARERLDPELTTLTGDSYGRAYDTVVRIQMLAELEEIAYKDHADEPARQEMQ 1489

Qy   876 RQIWWERLQGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVL 935
         | :: | |:||||:|::||  :|||::|||:|: || |:| |  |:|  :|| | ::|:|
Db  1490 RQTWKKRLAGCQRDVEVWQRILQVRSLVLVLKPNEDMDTWIEFADLCRTSDRLNLAEKTLTS 1549

Qy   936 LLGVD-PSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDA-------FQEMQHFVQTMQQQ 987
         ::|::   :|:  :|: | |:| :| ::  |:| ::| |       |:|:||| |:|:|
Db  1550 LVGFQYPS--MEDTGRAPPPIIFAYLRMAWAKNLQIDSREERYETLQHLRDFTDQLTDD 1607

Qy   988 A------QHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPK-VLQYYSAA 1039
         |       :||:||| |::|::  :|||| |:|||| ||:||:||::|:  ||||:::|
Db  1608 VGIGARGPNGRLMLPDQKLYGSYTKLLAQCHVELGQWQATLRESQSSADPSGILHDYCLA 1667

Qy  1040 TEHDRSWYKAWHAWAVMNPEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATAT 1099
         ||:| || | ||  ||  |:|:|:  :|:|  |:        |   :| |
Db  1668 TELDPEWYQAWETWALANPEVITQLEVSQQ-----------GLSPVHFFT-------- 1706

Qy  1100 TTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFPRSLSLRGNN 1159
                                          ::||    |::|:  | |:||:|:||
Db  1707 --------------------------------YIIPAVEGFLKSISLSPGNS 1726

Qy  1160 LQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQITWLQVIPQLIARIDTPRPLVGRLIHQ 1219
         ||||||:||:|:||| ||| |:||::|| | | ||:|: || |||| |:||::|| |:
Db  1727 LQDTLRLLTLWFTYGYSSGVTAAVSQGLPTVNIDVWLEVIPQLIARIQTPRQSIQQLIVQ 1786

Qy  1220 LLTDIGRYHPQALIYPLTVASKSTTARHNAANKLIKNMCEHSNTIVQQAMMVSEELIRV 1279
         ||:|||: |||:||||||||:||:| |||   :|| |:     |::| |:|:| |::|:
Db  1787 LLHDIGKAHPQLIYPLTVASKSTVAARRTVAQNITHKMREHSPKIVDQAELVSTELIRA 1846

Qy  1280 AILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLM 1339
         ||||||:|| | ::|||:||||  |||:|:||:|| |::  |||| ::| |:|| |||
Db  1847 AILWHEMWYDGLEEASKHYFGDHDIPGMLGVLEPLHEIVENGPQTLRETSFIQSFGHDLR 1906

Qy  1340 EAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELA 1399
         :|::|  |:: :|||:|  ::|: |||:|||| |:||  |  |:|:| || ||:|:::||
Db  1907 IARBHLKRYRITDQGTETQQAWDVYYSVFQRLGKQLKLLHVIELQYVSPKLMAVRDLDIA 1966
```

```
RT  "Mutation in ACTOR affects embryo development.";
RL  Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR  EMBL; AF178967; AAG43423.1; -.
DR  HSSP; P42345; 1FAP.
DR  GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR  InterPro; IPR008938; ARM.
DR  InterPro; IPR003151; FAT.
DR  InterPro; IPR003152; FATC.
DR  InterPro; IPR000403; PI3_PI4_kinase.
DR  InterPro; IPR009941; TPR-like.
DR  Pfam; PF02259; FAT; 1.
DR  Pfam; PF02260; FATC; 1.
DR  Pfam; PF00454; PI3_PI4_kinase; 1.
DR  SMART; SM00146; PI3Kc; 1.
DR  PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR  PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR  PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ  SEQUENCE  2481 AA;  279187 MW;  DA663EA9A9366F93 CRC64;

Query Match        42.2%;   Score 3970.5;  DB 10;  Length 2481;
Best Local Similarity  43.6%;  Pred. No. 5.9e-282;
Matches  844;  Conservative 333;  Mismatches 517;  Indels 241;  Gaps  36;

Qy    1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKD-PPDDNPGVINNVL   59
Db  662 LEISADNKCREESAKLLGCLVRNCERLILPYVAPVQKALVARLSGTGVNANNIVTGVL   721

Qy   60 ATIGELAQVSGLEMRKWVDELFPIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYR  119
Db  722 VTVGDLARVGGLAMRQYIPELMPLIVEALMDGAAVAKREVAVSTLGQVVQSTGYVVTPVK   781

Qy  120 KYPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSE  179
Db  782 EYPLLLGLLLKLLKGDLVWSTRREVLKVLGIMGALDPHVHK------RNQQSLSGSH    832

Qy  180 SKSSQDSSDYSTSEMLVNMGNLPL--------DEFYPAVSMVALMRIFRDQSLSHHHTM   230
Db  833 GEVPRGTGD--SGQPIPSIDELPVELRPSFATSEDYSTVAINSLMRILRDASLLSYEKR   890

Qy  231 VVCAITPIFKSLGLKCVQFLPQVMEPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPY  290
Db  891 VVRSLMIIFKSMGLGCVPYLPKVLPELFHTVRTSDENLKDFITWGLGTLVSIVRQHIRKY   950

Qy  291 MDEIVTIMREFWVMNT-------SIQSTILLLEQIVALGGEFKLIYLPQLIPHMLRV--   341
Db  951 LPELLLSIVSELWSSFTLPGPIRPSRGLPVIHLLEHLCLALNDERRTYLPVIIPCPIQVLG 1010

Qy  342 -------FMHDNSPGRIVSIKLLAAIQLFGANLDDYLHLLLPPIVKLFDAPFAPLPSR   392
Db 1011 DAERFNDYTTYVPD--------ILHTLEVFGGTLDEHMILLLPALIRLIFKV-DAPVAIR  1059

Qy  393 KAALEIVDRLTESLDFTDYASRIIHPIVATLD-QSPELRSTAMDTLSSLVFQLGKKYQIF   451
Db 1060 RDAIKPLTRVIPCVQVTGHISALVHHLKLVLDGKNDELRKDAVDALCCLAHALGEDFTIF  1119

Qy  452 IPMVNKVLVRHRINHQRYDVLICR--------IVKGYTLADEE-------EDPLIYQH   494
Db 1120 IRSIHKLLLKHRLRHKEPEBIHARMRRREPLIVA--TTATQQLSRRLPVEVIRDPVIENE  1177

Qy  495 RMLRSGQDALASGPVETGPMKKLHVSTINLQXAWGAAARRVSKDDWLEWLRRLSLELLKD   554
Db 1178 -----------IDPFEEGTDRNHQVNDGRLRTPAGEASQRSTKSDWEEWMRHFSIELLKE  1225

Qy  555 SSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRSIELALTSQDI-AEV   613
Db 1226 SPSIALRTCAKLAQLQPFVGELFAAGFVSCWAQLNESSOKQLVRSLEMAFSSNIPPEI   1285

Qy  614 TCTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYKELPFQKGPTP--   670
Db 1286 LATLLNLAEFMEHDEK-PLPI----DIRILLGALAEKCRVFAKALHYKEMEFE-GPRSKRM 1339

Qy  671 -----AILESLIS:NNKLQQPEAAAGVLEYAWKHFGELEIQATWYEKLHEWEDALVAYDK   725


Db 1340 DANPVAVVEALIHINNQLHQHEAAVGILTYAQQHL-DVQLKESWYEKLQRWDDLAKAYTL  1398
Qy  726 KMDTNKDDPELM----LGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARWAAAAAW   781
Db 1399 KA-SQTTNPHLVLEATLGQMRCLAAIARWEELNNLCKEYWSPAEPSARLEMAPWAAQAAW  1457

Qy  782 GLGQWDSMEEYTCMI-----------PRDTHDGA----FYRAVLALHQDLFSLAQQCI   824
Db 1458 NMGEWDQMAEYVSRLDDGDETKLRGLASPVSSGDGSNGTFFRAVLLVRRAKYDEAREYV  1517

Qy  825 DKARDLLDAELTMAGESYSRAYGAMVSCHMLSELEVICYKLVP----ERREIIRQ      877
Db 1518 ERARKCLATELAAIVLESYERAYSNMVRVQQLSELEEVIEYYTLPVGNTIAEERRALIRN 1577

Qy  878 IWWERLQGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLL   937
Db 1578 MWTQRICGSKRNVEVWQALLAVRALVLPPTEDVETWLKFASLCRKSGRISCQAKSTLLKLL 1637

Qy  938 GVDPSRQLDHPLPTVHPQVTYAYMGMW---KSARKIDAFQHMQHFVQTMQQ--QAQHAI   992
Db 1638 PFDPEVSPENMQYHGPPQVMLGYLKYQWSLGEERKRKKEAFTKLQILTRELSSVPHSQSDI 1697

Qy  993 ATEDQQHKQELHKLMARCFLKLGEWQLNL-QGINESTIPKVLQYSAATEHDRSWYKAWH  1051
Db 1698 LASMVSSKGANVPLLARVNLKLGTWQWALSSGLNDGSIQEIRDAFDKSTCYAPKWAKAWH 1757

Qy 1052 AWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATATTTASTEGSNSES 1111
Db 1758 TWALFNTAVMSHYISRGQIASQ------------                           1779

Qy 1112 EAESTENSPTPSPLQKKVTEDLSKILLMYTVPAVQGFFRSISLSRG-----NN-QDTLRV 1166
Db 1780 ------------YVVSAVTGVFYSIACAANAKGVDDSLQDILRL                 1811

Qy 1167 LTIWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGR 1226
Db 1812 LTILWFNHGATADVGTALKTGFSHVNINTWLVVLPQIIARIHSNNRAVRELIQSLLIRIGE 1871

Qy 1227 YHPQALLYPLTVASKSTTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHEM 1286
Db 1872 NHPQALMYPLLVACKSISNLRRARAAQEVVDKVRQHSGALVDQAQLVSHELIRVAILWHEM 1931

Qy 1287 WHEGLEEASRLYFGERNVKGWFEVLEPLHAMMERGPQ---TLKETSFNQAYGRDLMEAQ  1342
Db 1932 WHEALEEASRLYFGEHNIEGMLKVLEPLHDMLDEGVKKDSTTIQERARFIEAYRHELKEAH 1991

Qy 1343 EWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVSPKLLMCRDELAVPG   1402
Db 1992 ECCCNYKITGKDAELTQAWDLYYHVFKRIDKQLASLTTDLESVSPELLCRDELAVPG    2051

Qy 1403 TYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLF 1462
Db 2052 TYRADAPVVTISSFSRQLVVITSKQRPRKLTIHGNDGEDYAFLLKGHEDLRQDERVMQLF  2111

Qy 1463 GLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLN 1522
Db 2112 GLVNTLLENSRKTAEKDLSIQRYSVIPLSPNSGLIGWVPNCDTLHHLIREHRDARKIILN 2171

Qy 1523 IEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTPAGDDLAKLLWLKSPSSEVWFDRRTNYTR 1582
Db 2172 QENKHMLSFAPDYDNLPLIAKVEVFEYALENTEGNDLSRVLWLKSRSSEVWLERRTNVTR  2231

Qy 1583 SLAVMSMVGYIILGLGDRHPSNLMLDRLSGKLIHIDFGDCFEVAMTREKFPEKLPFRLTRM 1642
Db 2232 SLAVMSMVGYILGLGDRHPSNLMLHRYSGKLIHIDFGDCFEASMNREKFPEKVPFRLTRM  2291

Qy 1643 LTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTN------T 1696
Db 2292 LVKAMEVSGIEGNFRSTCENVMQVLRTNKDSVMAMMEARFVHDPLINWRLFNFNEVPQLAL 2351

Qy 1697 KGNKKSRTRTDSYSAGQSVEILDGVELGEPAHKKYGTTVPESIHSFIGDGLVKPEALNKK 1756
Db 2352 LGNNNPNAPADVEPDEDDEDPAD-IDLPQPQRSTREKEILQAV-NMLGDA--NEVLNER   2406
```

```
Qy     61  VNDETQAKMARMAAAAAWGLGWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQC  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    824  IDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSLEEVIQYKLVPERREIIRQIWWERL  883

Db    121  IDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSLEEVIQYKLVPERREIIRQIWWERL  180

Qy    884  QGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLIALAHKTLVLLLGVDPSR  943
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  QGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLIALAHKTLVLLLGVDPSR  240

Qy    944  QLDHPLPTVHPQVTTAYMKQMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQEL  1003
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  QLDHPLPTVHPQVTTAYMKQMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQEL  300

Qy   1004  HKLMARCFLKLGEWQLNLQGINESTIPKVLQYYSAATHDRSWYKAWHAWAVNFEAVLH  1063
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  HKLMARCFLKLGEWQLNLQGINESTIPKVLQYYSAATHDRSWYKAWHAWAVNFEAVLH  360

Qy   1064  YKHQNQARDEKKKLRHASGANITNATTAATTAATTAATTTASTEGSNSBEEAESTENSPTPS  1123
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  YKHQNQARDEKKKLRHASGANITNATTAATTAATTAATTTASTEGSNSESEAESTENSPTPS  420

Qy   1124  PLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVITLIWFDYGHWPDVNEAL  1183
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  PLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVITLIWFDYGHWPDVNEAL  480

Qy   1184  VEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKST  1243
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  VEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKST  540

Qy   1244  TTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERN  1303
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERN  600

Qy   1304  VKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDL  1363
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  VKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLNEAQEWCRKYMKSGNVKDLTQAWDL  660

Qy   1364  YYHVFRRISKQLPQLTSLELQVVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVI  1423
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  YYHVFRRISKQLPQLTSLELQVVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVI  720

Qy   1424  TSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQ  1483
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  TSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQ  780

Qy   1484  RYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQK  1543
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  RYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQK  840

Qy   1544  VEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGD  1598
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  VEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGD  895
```

```
Qy   1524  EHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEWFDRRTNYTRS  1583
               |||||:||: ::||||||||||||||||||||||||||||||:|||| |||||||||
Db   2072  EHRIMLQMCSDYDNLTLLQKEVEVQYALENSNGLDLHKVLWLKSRNSEVWLDRRTNVTRS  2131

Qy   1584  LAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLTRML  1643
               |||||||||||||||||||||||||||||||||||||||:|||||:|||||||||||||
Db   2132  LAVMSMVGYILGLGDRHPSNLMLDRHTGHILHIDFDGCFEVAMHRDKYPEKIPFRLTRML  2191

Qy   1644  TNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRSR  1703
               |||:|||||||||:||:||||||||||||||||||||||:||:||||||||||||||
Db   2192  INAMEVSGIEGNFRLTCEAVMNVLRNNKESLMAVLEAFVHDPLINWRLLTPNENNTKHKA  2251

Qy   1704  TR-TDSYSAGQSVEILDG-------VELGEPAHKKTGTTVPES-IH-------SFIG--  1744
               :: :   :   : :::                ::         ::
Db   2252  TNIASNNSTSNSTTKIEGDLNTIDNPINKESPDHEAVAGSLKSSPVHGRQIARNQRVGVD  2311

Qy   1745  ---DGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  1800
                  ::|::|::: ::||:|   : : :||||||||||||: ||||||||:|||||||||
Db   2312  AEQVEAEIVPEALNERALSVINRVNKKLTGRDFSSNEFLDVPBQVQKLIDQATSHENLCL  2371

Qy   1801  CYIGWCPFW  1809
               |:||||||
Db   2372  SYVGWCPFW  2380
```

```
Query Match             50.1%;  Score 4720;  DB 4;  Length 895;
Best Local Similarity   100.0%;
Matches  895;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    704  EIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTL  763
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  EIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTL  60

Qy    764  VNDETQAKMARMAAAAAWGLGWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQC  823
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    2438 KSVNEQSQVELLIQQAINNENLCQCYIGWCPFW 2470
```

RESULT 3
Q86C65

```
ID   Q86C65      PRELIMINARY;      PRT;    2380 AA.
AC   Q86C65;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tor.
GN   TOR.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=12626495;
RA   Otto G.P., Wu M.Y., Kazgan N., Anderson O.R., Kessin R.H.;
RT   "Macroautophagy is Required for Multicellular Development in the
RT   Social Amoeba Dictyostelium discoideum.";
RL   J. Biol. Chem. 278:17636-17645(2003).
DR   EMBL; AY204354; AAO43977.1; -.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000357; HEAT.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS50077; HEAT_REPEAT; 1.
DR   PROSITE; PS00915; PI3_4_KINASE 1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE 2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE 2380 AA; 268747 MW; 5E2998128B92711E CRC64;

Query Match   50.5%;  Score 4749.5;  DB 5;  Length 2380;
Best Local Similarity  51.7%;  Pred. No. 0;
Matches 967;  Conservative 290;  Mismatches 433;  Indels 179;  Gaps 31;
```

```
Qy      1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPIIKALILKLKDPDPDPNPGVINNVLA 60
            ||||   :  :|||||  :||:|:|  |  :||||||:|| ||||||||
Db    631 LEFSGDGRNKEESARLLGHLISASEKLIKPYVEPIIKALILPKLR----DSNPRVASCVLA 686

Qy     61 TIGELAQVSGLEMRKWVDELPIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
            |||  ||:|:||| :||:||  ||   :||  |  :|||:|:|||:|:::|||:||||||||
Db    687 ALGELSVVGGEEMVQHIDSLLPLIIDTLQDQSSTSKREVALKTLAQLASSTGYVVKPFSK 746

Qy    121 YPTLLEVLNFLKTEQNGQTRREAIRVLGLLGALDEYKHKVNIGMIDQSRDASAVSLSES 180
            |  |||:||:    :||||||:|||||||||||||:||:|||  :||||  |  :|:  :
Db    747 YPMLLDTLLNAIKTERIGSIRREVIKVLGILGSLDPYKHKMN--ELGKRRB------DP 797

Qy    181 KSSQDSSDYSTSEML-VNMGNLPLDEFYPAVSNVALMRIFRDQSLSHHHTMVVQAITFIF 239
            |::  :|:|::: :| ||:::|||:| |||||:|:::: : :|:: |||:|||::| ||:
Db    798 KANDDKNNNMTNIVITISPSN---EDYPTVALTALMKIILRDPSLSSHHTSVIQAVMYIF 854

Qy    240 KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREPLFQQLGMLVSFVKSHIRPYMDEIVTLMR 299
            ||||| | |:| :|| |:|:  :| ||  :||||| ||:|::|: || ||||||:||| |
Db    855 KSLSLKSIPPLPQIMPPFLHAMNTGEPLFREFLFQQLGSLVSIVKQHIRDYLVNVFALIE 914

Qy    300 EFWVMNTSIQSTILLIEQIVVALGGEEPKLYLPQLIPEMLRVFMHDNSPGRIVSIKLLAA 359
            :: |:  : |:|::   :  ::|:    :    |:  |:   ||:  :  | : |:|:||:
Db    915 KYW--NSNLLIPIIKLVEEISSALNDEFKVYLPNLIPQMLNVLHTDRSPFKRSPTTKVLRA 972

Qy    360 IQLFGANLDDYLHLLLPPIVCLFDAPEAPLPSRKAALETVDRLITESLDFTDYASRIIHPI 419
            :::|| :  |:|||::|| ::|| : |  :  :|| :    :|:::|:||| :|||:||||:|
Db    973 LEVFGTNLDDYLHLIVIPAIVKLFBQVDVTTQVRTLAIQTIGRICKKLNFSDYASRIIHPL 1032

Qy    420 VRTLDQS-PELRSTAMDTLSSLVFPJLGKKYQIFIPMWNKVLVRHRINHQRYDVLICRIVK 478
            |:|| :  ||  |: || :::|:  |  :|| |: |||| |::|  : :  ::|: :::
Db   2012 LVNNLLSANHETAKSHLSIRRFSVIPLSPNSGLIGWVPHSDTLHTLIKDFRDSNKILLSI 2071
```

```
Db   1033 ARVLDSTESBLREETLNTLCALVVQLGSDYAIFIPMVGKVLARREIQSTNYELLISKLLK 1092

Qy    479 GVTLADEEEDPLIYQHRMLRSGQGDALASG-----------------------PVETGPMK 516
Db   1093 N--------QQLMLTPSSGDDGMGANRFGGDHNGHHLGEDHNNTSTPLDIG-VK 1138

Qy    517 KLHVSTINLQKAWGAARRVSKDDWLEWLRRLSLELLKQSSSPSLRSCWALAQAYNPMARD 576
Db   1139 KLKANEQHLRNAWETSQRSTKEDWGEWIRRFSVELLRESSPSALRSCLSLAQDYHPLVKE 1198

Qy    577 LFNAAFVSCWSELNEDQQDELIRSIELALTSQDI-AEVTQTLLNLAEFMEHSDKGPLPLR 635
Db   1199 LFNAGFVSCWTELHEQFQEELVRSLETALLSPNIPPELLQTLLNLAEFMELHEK-PLPI- 1256

Qy    636 DDNGIVLLGERAAKCRAYAKALHYKELEFQKGPTPAILESLISINNKLQQPEAAAGVLEY 695
Db   1257 ---DIRTLGALABKCHAYAKALHYKESRFSQSPSSTI-EALISINNQLQQPEAAIGLLIY 1312

Qy    696 AMKHFGELBIQATWYEKLHEWEDALIVAYDKKMDTNKDDP------EIMLGMRCLBAIGEW 750
Db   1313 AQKNH-SVELKEGWYEKLRRWEDALIAAYEKK---QKDDPNGGTIENTMGIIRCLHALGEW 1368

Qy    751 GQLHQQCCBKW-TLVNDETQAKRMARMAAAAAWGLGQWDSMEBYTCMIPRDIHDGAFYRAV 809
Db   1369 ERLSALSSETWKSDINDHTRATIAPLAAAAAWNLVNWDKMDEYVCAWNKDTVEGSFYRAI 1426

Qy    810 LALHQDLFSLAQQCIDKARDLLDAELFAMAGESYSRAYGAMVSCHMLSELEEVIQYKLV- 868
Db   1427 LEVHHDNFTLAHGFIDHARTLVDTELTALLGESYNRAYKVVVRLQQLSELEEIIEYKKCV 1488

Qy    869 --PERREIIRQIWWERLQGSCQRIVEDWGKILMVRSLVVSPHEDMRTWLKYASLCGKSGRL 926
Db   1489 DSPERRNMIKNIWKTRLRGCCQHNVDIWQSILAVHSLVISPHEELDMWLKFVGLCRKGSRL 1548

Qy    927 ALAHKTLVLLGVDPS--RQLDHFLPTVHPQVTAYMKNMWKSARKIDAFQHMQHFVQTM 984
Db   1549 GLIAQKTLTMLMGKDPSTTSQFGSVLPNTHPRITFAYIKQLWSAGAQXPAFEKLRTFVQAL 1608

Qy    985 QQQAQHAIATEDQQHKQELHKLMARCFLKLGSWQLNL-QGINESTIPKVLQYYSAATEHD 1043
Db   1609 RD--------TDDL-------PLQGRAYIKLGEWQLALGDTLSEASIPHLISSFKAATBCD 1654

Qy   1044 RSWYKAWHAWAVMNFBAVLHYKHQNQARDEKKLRHASGANTINATTAANTAATATTTAS 1103
Db   1655 PNWYKAWHSWALLINFEVVSHYE-------- 1676

Qy   1104 TEGSNSESEAESTENSPTBSPLQKKVTEDLSKTLMYTVPAVQGFFRSISLSRGNNLQDT 1163
Db   1677 ---------EQIGAHLL-----PAVISFFKSISLGPDRSLQDT 1711

Qy   1164 LRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTD 1223
Db   1712 LRLLTLWFKHGAQKEVEAALMQGRNTIISIDTWLHVIPQIIARIHAPVLPVRLLHELIDT 1771

Qy   1224 IGRYHPQALIYPLTVASKTTTAPHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILW 1283
Db   1772 IGKEHPQALVYPLTVATKSHSPARLAAAKSLMDKRMRKEHSATLVDQALPVSQELVRTAILW 1831

Qy   1284 HEMWEHGLBEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE 1343
Db   1832 LEMWYEGLBEASRQYFGDHNPEAMLANTLAPLHQILHQLIELKGPETTSETSFLQAFGRDLQEALF 1891

Qy   1344 WCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGT 1403
Db   1892 WSKKYEKTRKEGDLNQAWDLYYQVFRRIYKQLPQMSSLELQYVSPKLLNSNMMELAVPGT 1951

Qy   1404 YDPNQPIIRIQSIAPSLQVITSKQRPKLITMGSSNGHEFVFLLKGHEDLRQDERVMQLFG 1463
Db   1952 VKASENVIRIQSFSQALSVIPSKQRPKRKLIIIGSDGLEYTFLLKGHEDLRQDERVMQLFS 2011

Qy   1464 LVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNI 1523
Db   2012 LVNNLLSANHETAKSHLSIRRFSVIPLSPNSGLIGWVPHSDTLHTLIKDFRDSNKILLSI 2071
```

```
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003638; AAF53337.1; -.
DR   HSSP; P42345; 1FAP.
DR   FlyBase; FBgn0021796; Tor.
DR   GO; GO:0030307; P:positive regulation of cell growth; NAS.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008940; Prenyl_trans.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE 2470 AA; 281031 MW; 5D78D2ECC07C7FF9 CRC64;

Query Match      57.3%; Score 5389.5; DB 5; Length 2470;
Best Local Similarity  58.5%; Pred. No. 0;
Matches 1073; Conservative 244; Mismatches 413; Indels 103; Gaps 17;

Qy    1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPIIKALILLKLKDPDPDNPGVINNVLA 60
Db  717 LKYSGMSRNKEQSAKMLDHLVISTPRLISSYMNPIIKALVPKLH--EPESNPGVILNVLR 774

Qy   61 TIGELAQVSG--LEMRKWTVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVBPY 118
Db  775 TIGDLAEVNGGSDEMLWADDLLSILLEMLGDAGSPDKRGVALWTLGQLISATGRVVTPY 834

Qy  119 RKYFTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLS 178
Db  835 HKYPVLIDILINFLKTEQRSIRRETIRVLGLLGAMDPYKHKMNKGLIDSQKDNVLIAYS 894

Qy  179 ESKSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFI 238
Db  895 DGK-VDESQDISTAELLVNMGN-ALDEYYPAVAIAALMRILRDPTLSTRHTSVVQANTFI 952

Qy  239 FKSLGLKCVQFLPQYMPTFLNVIRVCDGAIREFLPQQLGMLVSFVKSHIRPYMDEIVTLM 298
Db  953 FQSLGIKCVPYLAQVLPNLLDNVRTADNNLREFLFRQQLAILVAFVKLHIISYMGDIPKLI 1012

Qy  299 REFWVMNTSIQSTILLIEQIVALGEBFKLYLPQLIPHMLRVFMEDNSPGRITVSIKLLA 358
Db 1013 KEFWTINTPLQNTLINLINLIEQIAVALGCEFRDYLAELIPQILRVLQHDNSKDRMVTRLLQ 1072

Qy  359 AIQLFGANLDDYIHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHP 418
Db 1073 ALQFKFGSTLGYYLPLILPPIVKLFDSPYVPQQVSMVALETINNLACQLDFTDFSSRIIHP 1132

Qy  419 IVRTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVK 478
Db 1133 LVRVLDARPELRDQAMTTLRSLAKQLG3KKYLVFFVMVQRTLNKHRIVDPEYEELLSKIKS 1192

Qy  479 GYTLADEEEDPLIYQHRMLRSGCQGDALASGPVETGPM----------KKLHVSTINLQKAWG 530
Db 1193 CSTLADS---------YGAGRSELRPSRPKNNEPFVTDRNSNNKNLQVTNELRTAWQ 1241

Qy  531 AARRVSKDDWLEWLRRLSLELLKDSSSFSLRSCWALAQAYNPMARDLFNAAFVSCWSELN 590
Db 1242 VTRRVSKDDWVEWLKRLSIGLLKESPSHALRACRSLAQEYDTLLRDLFNAAFISCWTELS 1301

Qy  591 EDQQDELIRSIELALTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKC 650
Db 1302 PDLKNELTQSLIQALQVTDMPEITQTILNLAEFMEHCDRDPIPIETK----LLGTRAMAC 1357

Qy  651 RAYAKALHYKELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKGFGELEIQATWY 710
Db 1358 RAYAKALRYKEEFILLREDSQVFESLLLINNKLQQREAAGLLTRYRNAANELNVQGRWY 1417

Qy  711 EKLHEWEDALVAYDKKMDTNKDDPEELMLGRMRCLEALGEWGQLHQQCCKKWTIVNDETQA 770
Db 1418 EKLHNWDEALEHYERNLKTDSSDLEARLGEMRCLEALGDWSELSNVTKHEWENFGTEAKS 1477

Qy  771 KMARMAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDL 830
Db 1478 RAGPLAAVAAWGLQDWEAMREYVRCIPEDTQDGSYRAVLAVHHDFETAQRLLDETRDL 1537

Qy  831 LDAELTAMAGESYRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIV 890
Db 1538 LDTELTSMAGESYERAYGAMVCVQMLAELEEVIQYKLIPERREPLKTMWWKRLQGGQRLV 1597

Qy  891 EDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLP 950
Db 1598 EDWRRIIQVHSLVVKPHEDIHTWLKYASLCRKSGSLHLSHKTLVMLLGTDPKLNPNQPLP 1657

Qy  951 TVHPQVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQAQHAIATEDQQHKQELHKLMARC 1010
Db 1658 CNQPQVTYAYTKYMAANNLQEAYEQLTHFVSTYSQELS---CLPPEALKQQDQRLMARC 1714

Qy 1011 FLKLGEWQLNLQ-GINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKIQNQ 1069
Db 1715 YLRMATWQNKLQDSIRPDAIQGALECFEKATSYDPNWYKAWHLWAYMNFKVV---QAQKS 1771

Qy 1070 ARDEKKKLRHASGANITNATTAATAATATTTASTEGSNSESBAESTENSPTPSPLQKKV 1129
Db 1772 ALDKQQ----PPGASM-----------------------GNTMGGSGLDSDL------- 1795

Qy 1130 TEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKA 1189
Db 1796 -----MIQRYAVPAVQGFFRSISLIKGNSLQDTLRLLTLWFDYGGNHABVYEALLSGMKL 1850

Qy 1190 IQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHN 1249
Db 1851 IEINTWLQVIPQLIARIDTHRQLVGQLIHQLLMDIGKNHPQALVYPLTVASKSASLARRN 1910

Qy 1250 AANKILKNMCEHSNTIVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFE 1309
Db 1911 AAFKILDSMRKHSPTIVEQAVMCSEELIRVAILWHEQWHEGLEEASRLYFGDRNVKGMFE 1970

Qy 1310 VLEPLHAMMERGPQTIKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFR 1369
Db 1971 ILEPIHAMLERGPQTKETSFSQAYGRELTBAYEWSQRYKTSAVVMDLDRAWDIYYHVFQ 2030

Qy 1370 RISKQLPQLTSLELQYVSPKLLMCRDIELAVPGTYDPNQPIIRIQSIAPSLQVTTSKQRP 1429
Db 2031 KISRQLPQLTSLELPYVSPKLMTCKDLELAVPGSYNPSQELIRISIIKTNLQVITSKQRP 2090

Qy 1430 RKLTLMGSNGHBFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKGNLSIQRYAVIP 1489
Db 2091 RKLCIRGSNGKDYMYLLKGHEDLRQDERVMQLFSLVNTLLLDDPTFRRNLAIQRYAVIP 2150

Qy 1490 LSTNSGLIGWPHCDTLHALIRDYREKKKLLLANIEHRIMLRMAPDYDHLTLMQKVEVFEH 1549
Db 2151 LSTNSGLIGWPHCDTLHTLIRDYRDKKVPLNQEHRIMLNFAPDYDHLTLMQKVEVFEH 2210

Qy 1550 AVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRL 1609
Db 2211 ALGCTQGDDIAKLLWLKSPSSELWFERRNNYTRSLAVMSMVGYILGLGDRHPSNLMLDRM 2270

Qy 1610 SGKILHIDFGDCFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLRE 1669
Db 2271 SGKILHIDFGDCFEVAMTREKFPEKIPFRLTRMLIKAMEVTGIEGTYRRTCRSVMLVLRR 2330

Qy 1670 HKQSVMAVLEAFVYDPLLNWRLMDYNTKGNKSRSRTRTDSYS---------AGQSVEILD 1719
Db 2331 NKDSLMAVLEAFVYDPLLNWRLLDVDKKGN------DAVAGAPGGRGGSGMQDSLSN 2383

Qy 1720 GVELGEPAHKK---TGTTVPESIHSFIGDGLVKPEALNKKAIQINRVRDKLTGRDFSHD 1776
Db 2384 SVEDSLFMAKSKPYDPTLQQGLHHNVAD-----ETNSKASQVIKRVKCKLJTGTDFQTE 2437

Qy 1777 DTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
```

```
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS52290; PI3_4_KINASE_3; 1.
FT   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1188 AA;  136108 MW;  OECEDCA1D9F163CB CRC64;

Query Match            66.6%;  Score 6265;  DB 4;  Length 1188;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1188;          Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   622 EFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYKELEFQKGPTPAILESLISINN 681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 EFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYKELEFQKGPTPAILESLISINN 60

Qy   682 KLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRM 741
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 KLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRM 120

Qy   742 RCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAWGLGWDSMEEYTCMIPRDTH 801
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 RCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAWGLGWDSMEEYTCMIPRDTH 180

Qy   802 DGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEE 861
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 DGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEE 240

Qy   862 VIQYKLVPERREIIRQIWWERLQGCCRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCG 921
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 VIQYKLVPERREIIRQIWWERLQGCCRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCG 300

Qy   922 KSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQHMQHFV 981
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 KSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQHMQHFV 360

Qy   982 QTMQQQAQHAIATEDQOHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQYYSAATE 1041
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 QTMQQQAQHAIATEDQOQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQYYSAATE 420

Qy  1042 HDRSWYKAWEAWAVMNFEAVLYHKHQNQARDEKKKLRHASGANITNATTAATTAATATTT 1101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 HDRSWYKAWEAWAVMNFEAVLYHKHQNQARDEKKKLRHASGANITNATTAATTAATATTT 480

Qy  1102 ASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLMYTVPAVQGFFRSISLSRGNNLQ 1161
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLMYTVPAVQGFFRSISLSRGNNLQ 540

Qy  1162 DTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLL 1221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 DTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLL 600

Qy  1222 TDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTIVQQAMMVSEELIRVAI 1281
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTIVQQAMMVSEELIRVAI 660

Qy  1282 LWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTIKETSFNQAYGRDLMEA 1341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 LWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTIKETSFNQAYGRDLMEA 720

Qy  1342 QEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVP 1401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 QEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVP 780

Qy  1402 GTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQL 1461
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQL 840

Qy  1462 FGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYREKKKLLL 1521
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 FGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYREKKKLLL 900

Qy  1522 NIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFPRRTNYT 1581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 NIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFPRRTNYT 960

Qy  1582 RSLAVMSMVGYILGLGDRHPSNLM.DRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLTR 1641
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 RSLAVMSMVGYILGLGDRHPSNLM.DRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLTR 1020

Qy  1642 MLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKR 1701
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 MLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKR 1080

Qy  1702 SRTRTDSYSAGQSVBILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQII 1761
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 SRTRTDSYSAGQSVBILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQII 1140

Qy  1762 NRVRDKLTGRDFSHDDTLDVPFQVELLIKQATSHENLCQCYIGWCPFW 1809
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 NRVRDKLTGRDFSEDDTLDVPFQVELLIKQATSHENLCQCYIGWCPFW 1188
```

```
RESULT 2
Q9VK45
ID   Q9VK45          PRELIMINARY;      PRT;   2470 AA.
AC   Q9VK45;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG5092 protein.
GN   TOR OR CG5092.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Murphy D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2004, 19:58:43 ; Search time 62 Seconds
(without alignments)
9206.007 Million cell updates/sec

Title: US-09-517-491-12
Perfect score: 9413
Sequence: 1 LEHSGIGRIKEQSARMLGHL........KQATSHENLCQCYIGWCPFW 1809

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6265	66.6	1188	4	Q96Q3	Q96qg3 homo sapien
2	5389.5	57.3	2470	5	Q9VK45	Q9vk45 drosophila
3	4749.5	50.5	2380	5	Q86C65	Q86c65 dictyosteli
4	4720	50.1	895	4	Q96QW8	Q96qw8 homo sapien
5	3970.5	42.2	2481	10	Q9FR53	Q9fr53 arabidopsis
6	3852	40.9	2360	3	094188	094188 cryptococcu
7	3851	40.9	2513	10	Q9LPM4	Q9lpm4 arabidopsis
8	3427.5	36.4	2692	5	Q95Q95	Q95q95 caenorhabdi
9	3425	36.4	2695	5	O01438	O01438 caenorhabdi
10	2703	28.7	1083	5	094189	094189 cryptococcu
11	2646.5	28.1	760	5	Q960H0	Q960h0 drosophila
12	2064	21.9	2604	5	Q816M7	Q816m7 giardia lam
13	2036.5	21.6	1583	5	076222	076222 trypanosoma
14	1667	17.7	455	5	Q8T916	Q8t916 drosophila
15	1621	17.2	651	3	Q9HFM9	Q9hfm9 glomus moss
16	1142	12.1	2903	5	Q9N8R7	Q9n8r7 trypanosoma
17	1101.5	11.7	374	10	Q8RWA9	Q8rwa9 arabidopsis
18	955	10.1	249	13	Q9YI30	Q9yi30 carassius a
19	932	9.9	180	11	Q9CST1	Q9cst1 mus musculu
20	889	9.4	2454	3	Q9UVP2	Q9uvp2 emericella
21	884	9.4	2454	3	Q9UV56	Q9uv56 emericella
22	850	9.0	280	10	Q9FKS4	Q9fks4 arabidopsis
23	848.5	9.0	2703	10	Q9MAZ4	Q9maz4 arabidopsis
24	807	8.6	3218	5	Q9W3V6	Q9w3v6 drosophila
25	753.5	8.0	2644	4	Q13535	Q13535 homo sapien
26	750	8.0	3454	10	Q7Y1G0	Q7y1g0 oryza sativ
27	745.5	7.9	2654	13	Q9DDK7	Q9ddk7 xenopus lae
28	737	7.8	2344	5	Q86H67	Q86h67 dictyosteli
29	734.5	7.8	2654	13	Q9DE14	Q9de14 xenopus lae
30	729	7.7	3521	4	Q8NFX2	Q8nfx2 homo sapien
31	726	7.7	3657	4	Q96Q15	Q96q15 homo sapien
32	693	7.4	3031	4	Q96QV0	Q96qv0 homo sapien
33	683.5	7.3	2392	4	Q96RW3	Q96rw3 homo sapien
34	663.5	7.0	2354	5	Q24135	Q24135 drosophila
35	663.5	7.0	2517	5	Q9VXG8	Q9vxg8 drosophila
36	662	7.0	1988	4	O43305	O43305 homo sapien
37	645	6.9	2322	5	O01510	O01510 caenorhabdi
38	641.5	6.8	3856	10	Q9M4D7	Q9m4d7 arabidopsis
39	640.5	6.8	3856	10	Q9M3G7	Q9m3g7 arabidopsis
40	619.5	6.6	2429	5	Q9VFB1	Q9vfb1 drosophila
41	605	6.4	1579	13	Q9PTMO	Q9ptm0 xenopus lae
42	585.5	6.2	2812	3	O74630	O74630 schizosacch
43	564.5	6.0	4133	13	Q9DEI2	Q9dei2 gallus gall
44	564.5	6.0	4134	13	Q80GX4	Q80gx4 gallus gall
45	564	6.0	4128	4	Q7Z6I1	Q7z6i1 homo sapien

ALIGNMENTS

RESULT 1
Q96QG3
ID   Q96QG3   PRELIMINARY;   PRT; 1188 AA.
AC   Q96QG3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Rapamycin associated protein FRAP2 (Fragment).
GN   FRAP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21318995; PubMed=11426320;
RA   Stover C., Endo Y., Takahashi M., Lynch N., Constantinescu C.,
RA   Vorup-Jensen T., Thiel S., Friedl H., Hankeln T., Hall R., Gregory S.,
RA   Fujita T., Schwaeble W.;
RT   "The human gene for mannan-binding lectin-associated serine protease-2
RT   (MASP-2), the effector component of the lectin route of complement
RT   activation, is part of a tightly linked gene cluster on chromosome
RT   1p36.2-3.";
RL   Genes Immun. 2:119-127(2001).
DR   EMBL; AJ300188; CAC15570.1; -.
DR   GO; GO:0004428; F:inositol/phosphatidylinositol kinase activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.

```
Db      1 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH    60

Qy   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV  1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV   120

Qy   1501 PHCDTLHALIRDYREKKKILNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA  1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 PHCDTLHALIRDYREKKKILNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA   180

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD   240

Qy   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA   300

Qy   1681 FVYDPLLNW  1689
          |||||||||
Db    301 FVYDPLLNW   309
```

Search completed: March 2, 2004, 20:01:24
Job time : 77 secs

DE   Candida RAPT1-like protein.
XX
KW   RAPT1; rapamycin binding protein; FKBP; immunosuppressive; fungicide;
KW   anti-mycotic; agonist; antagonist; cell proliferation.
XX
OS   Candida albicans.
XX
FH   Key   Location/Qualifiers
FT   Domain   41..173
FT   /label= Rapamycin-binding_domain
XX
PN   WO9533052-A1.
XX
PD   07-DEC-1995.
XX
PF   30-MAY-1995; 95WO-US006722.
XX
PR   27-MAY-1994; 94US-00250795.
PR   20-DEC-1994; 94US-00360144.
XX
PA   (MITO-) MITOTIX INC.
XX
PI   Berlin V, Chiu MI, Cottarel G, Damagnez V;
XX
DR   WPI; 1996-030563/03.
DR   N-PSDB; AAT33873.
XX
PT   Rapamycin binding protein RAPT-1 - used as agonist, or antagonist of
PT   rapamycin cellular proliferation regulation.
XX
PS   Claim 13; Page 92-95; 121pp; English.
XX
CC   Candida albicans RAPT1-like protein (AAW04236) is the product of a cDNA
CC   clone (AAT33873) isolated from a Candida library using a probe generated
CC   with PCR primers based on conserved regions of RAPT1 and TOR proteins.
CC   The rapamycin-binding domain of this and other rapamycin-binding proteins
CC   (see also AAW04234-38) can be used in drug screening assays to identify
CC   agents that modulate the binding of rapamycin binding proteins with FK506
CC   -binding proteins. Such agents can be used to alter the growth and/or
CC   differentiation of a cell or as cell culture additives
XX
SQ   Sequence 562 AA;

Query Match   18.4%; Score 1735; DB 2; Length 562;
Best Local Similarity   55.8%; Pred. No. 2e-143;
Matches 330; Conservative   90; Mismatches 129; Indels   42; Gaps   4;

Qy   1232 LIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMVSEBLIRVAILWHEMWHEGL 1291
          |:|||||| | |||:|| |||| |:|| ||:|: ::|| ||:||||| || |
Db   1 LVYPLTVAITSESTSRKKAAQSIIEKXRVHSPSLVDQAELVSRELIRVAVLWHBQWHDAL   60

Qy   1292 EEASRLYFGERNVKGMFEVLEPLHAMMERGPQTILKETISFNQAYGRDLMEAQEWCRKYMKS 1351
          |:||| | |||| || ||||||||| |:||:: |||:|||: ||| :|:|||:|||||
Db   61 EDASRFFFGEHNTEKMFETLEPLHQMLQKGPETMREQAFANAFGRELTDAYEBWVLNFRRT 120

Qy   1352 GNVKDLTQAWDLYFHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPII 1411
          :: ||| |||||:| || :|| :|||::| |||| |||||||:|||||||||| |||
Db   121 KDITNLNQAWDIYFNVFRRVSKQVQLLASLELQYVSPDLEHAQDLELAVPGTYQAGKPVI 180

Qy   1412 RTQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLIAN 1471
          | |||||||||||||||||||| ||||| || |||||||||| |||:||||||||||| |
Db   181 RIIKPDTFSIISSKQRPRKLSCRGSDGKDYQYALKGHEDIRQDNLVMQLFGLVNTLLVN 240

Qy   1472 DPPSLRKNLSIQRYAVIPLSTNSGLIGWTPHCDTLHALIRDYREKKKILLNIEHRIMLRM 1531
          || || ||||| || |||| ||:|||||||| || |:||| |||| || :|||: |:
Db   241 DPVCFKRHLDIQQYPAIPLSPKVGLLGWTVPNSDTFHVLIKGYRESRSIMLNIEHRLLLQM 300

Qy   1532 APYDHFTLMQKVEVFEHAVNNTAGDDLAKLLWLJKSPSSEVWFDRRTNYTRSLAVMSMVG 1591
          ||||: | ||| | ||:| ||| | |: | ||:| ||| |:|||||||||||
Db   301 APYDFLTLLQKVFFTSAMDNCKGQDLYKVLWLMKSKSSEAWLDRRTYTTRSLAVMSMVG 360

Qy   1592 YILGLGDRHPSNLMLDRISGKILHIDFGDCFEVAMTREKPPEKIPFRLTRMLTNAMEVTG 1651
          |||||||||||||||||| |||:|||||||| |:   :|| | ||| || || | ||:|:

```
Db   2246  DCFEASWNREKFPEKVPFRLTRMLVKAMEVSGIEGNFRSTCENVMQVLRTNKDSVWAMME 2305
           ||||::|  |||||:|||||||||     ||||:|:|| :||||||  ||||||::|
Qy   1680  AFVYDPLLNWRLMDTN------TKGNKRSRTRTDSYSAGQSVEILDGVELGBPAHKKTGT 1733
           ||:||||:||| ||  : |  || ||  | : :|:|
Db   2306  AFVHDPLINWRLFNFNEVPQLALLGNNNPNAPADVEPEEDBDPAD-IDLPQPQRSTREK 2364

Qy   1734  TVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFS---------HDDT---- 1778
           : ::: :|:  :  ::  ||  :: :|: |: |||||||       :::
Db   2365  EILQAV-NMLGDA--NEVLNRRAVVVMARMSHKLTGRDFSSSAIPSNPIADHNNLLGGD 2420

Qy   1779  ------LDVPTQYELLIKQATSHENLCQCYIGWCP 1807
                 :|:  |||||||||||||| :|:|
Db   2421  SHVEHGLSVKVQVQKLINQATSHENLCQNYVGVVP 2456
```

RESULT 13
ADE62172
ID   ADE62172 standard; protein; 1417 AA.
XX
AC   ADE62172;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human Protein CAB44736, SEQ ID NO 8101.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO200301675-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO) GEN HOSPITAL CORP.
PA   (FARB) BAYER AG.
XX
PI   Woolf C, D'urso D, Befort K, Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; CAB44736.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1417 AA;

Query Match   36.5%; Score 3440; DB 7; Length 1417;
Best Local Similarity   100.0%; Pred. No. 2.6e-293;
Matches   677; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1  LEHSGIGRIKEQSARMLGHIVSNAPRLIRPYMEPILKALILKLKDPPDDNPGVINNVLA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741  LEHSGIGRIKEQSARMLGHIVSNAPRLIRPYMEPILKALILKLKDPPDDNPGVINNVLA 800

Qy    61  TIGELAQVSGLEMRKWVDELFPIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801  TIGELAQVSGLEMRKWVDELFPIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 860

Qy   121  YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHHKVNIGMIDQSRDASAVSLSES 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861  YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHHKVNIGMIDQSRDASAVSLSES 920

Qy   181  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITPIFK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSVVALMRIFRDQSLSHHHTMVVQAITPIFK 980

Qy   241  SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981  SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy   301  FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMDNSPGRIVSIKLLAAI 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  1041  FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMDNSPGRIVS-KLLAAI 1100

Qy   361  QLFGANLDDYLHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1101  QLFGANLDDYLHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy   421  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRERRINHQRYDVLICRIVKGY 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1161  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRERRINEQRYDVLICRIVKGY 1220

Qy   481  TIADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQXAWGAARRVSKDDW 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1221  TIADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQXAWGAARRVSKDDW 1280

Qy   541  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1281  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy   601  IELALTSQDIABVTQTLLNLABFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1341  IELALTSQDIABVTQTLLNLABFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy   661  ELEFQKGPTPAILESLI 677
          |||||||||||||||||
Db  1401  ELEFQKGPTPAILESLI 1417
```

RESULT 14
AAW04236
ID   AAW04236 standard; protein; 562 AA.
XX
AC   AAW04236;
XX
DT   24-NOV-1996 (first entry)
XX

```
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K, Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms.
XX
PS   Claim 5; SEQ ID NO 514; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins (ABB90790-ABB94016)
CC   for herbicidally active compounds, comprising aligning and comparing
CC   nucleic acid or amino acid sequences from plant with nucleic acid or
CC   amino acid sequences from non-plant organisms using suitable search
CC   parameters, where plant sequences having an E-value greater by a factor
CC   of 3 than the E-value of most similar non-plant sequences are selected.
CC   The polypeptides or nucleic acids encoding them are useful for
CC   identifying modulators. The identified modulators are useful as
CC   herbicides
XX
SQ   Sequence 2467 AA;

Query Match             39.9%;  Score 3755.5;  DB 5;  Length 2467;
Best Local Similarity   41.7%;  Pred. No. 0;
Matches   816;  Conservative  327;  Mismatches  522;  Indels  291;  Gaps  37;

Qy     1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKD-PDPDNPGVINNVL  59
Db   643  LELSADNKCREBSAKLLGCLVRNCERIIILPYVAPVQKALVARLSEGTGVNANNIVTGVL  702

Qy    60  ATIGELAQVSGLEMRKWVDELFIIIMDMLQDDSLILAKRQVALWTLGQLVASTGYVVEPYR  119
Db   703  VTVGDIARVGGLAMRQYIPELMPLIVEALMDGAAVAKREVAVSTLGQVVQSTGYVVTPYK  762

Qy   120  KYPTLLEVLLANFLKTEQNGQTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSE  179
Db   763  EYPLLLGLLLKLLKGDLVWSTRREVLKVLGIMGALDPHVHK-------RNQQSLSGSH  813

Qy   180  SKSSQDSDSDYSTSEMLVNMGNLPL---------DEFYPAVSMVALMRIFRDQSLSHHHTM  230
Db   814  GEVPRGTGD--SGQPIPSIDELPVELRPSFATSEDYYSTVAINSLMRILRDASLLSYHKR  871

Qy   231  VVQAITFIFKSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPY  290
Db   872  VVRSLMIIFKSMGLGCVPYLPKVLPELFHTVRTSDENLKDFITWGLGSTLVSIVRQHIRKY  931

Qy   291  MDEIVTLMREFWVMNT----------------SIQSTIILLIEQIVVALGGEFKL  329
Db   932  LPEELSLVSELWSSFTLPCGPIPSRGLPISYGFYFASLTDNQVLHLEHICLALNDEFRT  991

Qy   330  YLPQLIPHMLRV----------FMHDNSPGRIVSIKLLAAIQLRGANLDDYLHLLLPPI  378
Db   992  YLPVILPCFIQVLGDAERFNDYTYVPD----------ILHTLEVFGGTLDEHVHLLLPAL  1041

Qy   379  VKLRFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIVRTLD-QSPELRSTAMDTL  437
Db  1042  IRLFKV-DAPVAIRRDALKTLTRVIPCVQVTGHISALVHHLKLVLDGKNDELRKDAVDAL  1100

Qy   438  SSLVFQLGKCKYQIFIIPMVNKVLVRHRINHQRYDVLICR--------IVKGYTLADEE----  486
Db  1101  CCLAHALGEDFTIFIESIHKLLLKHLRHLRHKEFEEIHARWRRREPLIVA--TTATQOLSRR  1158

Qy   487  ------EDPLIYQHRMLRSGGQDALASGPVETGPMKKLLHVSTINLQKAWGAARRVSKDDW  540
Db  1159  LPVFVIRDPVIENE----------IDPFEBGTDRNHQVNDGRLRTAGEASQRSTKEDW  1206

Qy   541  LEWIRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  600
Db  1207  EEWMREHPSIELLKESPSPALRTCAKLAQLQPFVGRELFAAGFVSCWAQLNESSQKQLVRS  1266

Qy   601  IELALTSQDI-AEVTQTLLNLAEFMEHSDKGPLPLRDNGIVLLGERAAKCRAVAKALHY  659
Db  1267  LEMAFSSPNIPPEILATLLNLAEFMEHDEK-PLPI----DIRLLGALAEKCRVPAKALHY  1321

Qy   660  KELEFQKGPTP---------AILESLISINNKLQQPEAAAGVLEYAMKHFG-------E  702
Db  1322  KEMEFE-GPRSKRMDANPVAVVEALIHINNQLHQHEAAVGILTYAQQHLDVQLKESCGKA  1380

Qy   703  LEIQATWYEKLHEWEDALVAYDKKMDTNKDPEIM----LGRMRCLEALGEWGQLHQQCC  758
Db  1381  LTVFVSRYEKLQRWDDALKAYTLKA-SQTTNPHLVLEATLGQMRCLAALARWEELNNLCK  1439

Qy   759  EKWTLVNDETQAKARMARAAAAWGLGQWDSMEEYTCMI----------PRDTHDGA-  804
Db  1440  EYWSPAEPSARLEMAPMAAQAAWNMGEWDQMAFRVSRLDDGDETKLRGLASPVSSGDGSS  1499

Qy   805  ---FYRAVIALHQDLFSLAQQCIDKARDLLDAELTAWAGESYRSAYGANVSCHMLSELEE  861
Db  1500  NGTFFRAVLLVRRAYDEAREYVERARKCLATELAALVLESYERAYSNVRVQQLSRLEE  1559

Qy   862  VIQYKLVP------ERREIIRQIWERLGQCQRIVEDWQKILMVRSLVVSPHEDMRTWL  914
Db  1560  VIEYYTLPVGNTIAEERRALIRNMWTQRIQGSKRNVEVWQALLAVRALVLPPTEDVETWL  1619

Qy   915  KYASLCGKSGRLALAHKTIVLLGVDPSRQLDHPLPTVHPQVTYAYMKNMW--KSARKI  971
Db  1620  KFASLCRKASGRISOAKSTLLKLLPFDDEVSPENMQYHGPPQVMLGYLKYQWSLGEERKRK  1679

Qy   972  DAFQHMQHFVQTMQQ--QAQHAIATEDQQHKQELHKLMARCFIKLGEWQLNL-QGINEST  1028
Db  1680  EAFTKLQILTRELSSVPHSQSDILASMVSSKGANVPLLARVNLKLGTWQWALSSGLNDGS  1739

Qy  1029  IPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKLRHASGANITNA  1088
Db  1740  IQEIRDAFDKSTCYAPKWAKAWHTWALFNTAVMSHYISRGQIASQ---  1784

Qy  1089  TTAATTAATATTTASTEGSNSEEASTENSPTPSPLQKKVTEDLSKTLLMTVPAVQGF  1148
Db  1785  ----------------------------YVVSAVTGY  1793

Qy  1149  FRSISLSRG-------NNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLI  1203
Db  1794  FYSIACAANAKGVDDSLQDILRLLTLWFNHGATADVQTALKTGFSHVNINTWLVVLPQII  1853

Qy  1204  ARIDTPRPLVGRLLIHQLLTDIGRYHPQALLYPLTVASKSTTARHNAANKILKNMCEHSN  1263
Db  1854  ARIHSNNRAVRELIQSLLIRIGENHPQALMYPLLVACKSISNLRRAAAQEVVDKVRQHSG  1913

Qy  1264  TLVQQAMMVSDELIRVAILWHEMWHEGLEEASRLYFGERNVKGMPEVLEPLHAMMERGPQ  1323
Db  1914  ALVDQAQIVSHELIRVAIIWHEMWHEALEEASRLYFGEHNIEGMKVLEPLHDMLDEGVK  1973

Qy  1324  ----TLKETSFNQAYGRDLMRAQEWCRKYMKSGNVKDLTQAWDLYYHVRRISKQLPQLT  1379
Db  1974  KDSTTIQERAFIEAYRHELKEAHECCCNYKITGKDAELTQAWDLYHVFKRIDKQLASLT  2033

Qy  1380  SLELQYYSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKITLMGSNG  1439
Db  2034  TLDIESVSPELLLCRDLLAVPGTYRADAPVVTISSFSRQLVVITSKQRPRKLTIHGNDG  2093

Qy  1440  HEFVFLLKGHEDLRQGDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGW  1499
Db  2094  EDYAFLLKGHEDLRQDERVMQLFGLVNTLLENSRKTAEKDLSIQRYSVIPLSPNSGLIGW  2153

Qy  1500  VPHCDTLHALIRDYREKKKLLHVSTINLQKAWGAARRVSKDDW  1559
Db  2154  VPNCDTLHHLIREHRDARKIILNQENKHMLSFAPDYDNLPLIAKVEVFEYALENTEGNDL  2213

Qy  1560  AKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFG  1619
Db  2214  ----------SRVGYILGLGDRHPSNLMLERYSGKILHIDFG  2245

Qy  1620  DCFEVAMTREKFPEKIPFRLTRMLTNAMEVTLGLDGNYRTTCHTVMEVLREHKDSVMAVLE  1679
```

```
FT             restore the number of residues to that stated in the
FT             specification (residues omitted by an apparent printing
               error)"
XX
PN   US2003032775-A1.
XX
PD   13-FEB-2003.
XX
PF   13-SEP-2001; 2001US-00950634.
XX
PR   08-MAR-1994; 94US-00207975.
PR   26-SEP-1994; 94US-00312023.
PR   13-FEB-1995; 95US-00384524.
PR   06-JUN-1995; 95US-0471112.
XX
PA   (AMHP ) AMERICAN HOME PROD CORP.
XX
PI   Caggiano TJ, Chen Y, Failli AA, Molnar-Kimber KL, Nakanishi K;
XX
DR   WPI; 2003-492083/46.
DR   N-PSDB; ACA63013.
XX
PT   Novel effector protein of rapamycin of mammalian origin for identifying
PT   immunomodulatory or anti-tumor agent, binds glutathione-S-transferase-FK-
PT   506 binding protein-rapamycin complex.
XX
PS   Disclosure; Page 21-24; 32pp; English.
XX
CC   The invention relates to an effector protein of rapamycin of mammalian
CC   origin having a molecular weight of about 125, 149, 208 or 210 kDa which
CC   binds to a glutathione-S-transferase (GST)-FK-506 binding protein (FKBP)-
CC   rapamycin complex. Also included are a recombinantly produced protein
CC   comprising a partial amino acid sequence of one of the proteins listed
CC   above, cDNA/DNA of mammalian origin which encodes the proteins and
CC   antisense DNA or RNA derived from cDNA/DNA clone. The proteins are useful
CC   for identifying an immunomodulatory or anti-tumour agent, for detecting,
CC   in a biological sample, rapamycin, rapamycin analogues or metabolites.
CC   Antisense RNA or DNA is useful for modulating the immune system of a
CC   mammal in need of it. The proteins are useful for identifying, designing
CC   or synthesising immunomodulatory, anti-restenosis or anti-tumour agents,
CC   in screening assays, such as enzyme inhibitor assays and binding assays,
CC   to identify endogenous complexes and ligands and novel exogenous
CC   compounds (e.g. rapamycin) which modulate their functions, and in assays
CC   to identify a compound with therapeutic benefit for restenosis
CC   immunomodulation and as antitumour agents. Antisense DNA or RNA is useful
CC   for treating transplantation rejection, graft versus host disease,
CC   autoimmune diseases such as lupus, myasthenia gravis, multiple sclerosis,
CC   rheumatoid arthritis, type I diabetes, and diseases of inflammation such
CC   as psoriasis, dermatitis, eczema, seborrhea, inflammatory bowel disease,
CC   pulmonary inflammation, asthma, and eye uveitis, as an immunosuppressive
CC   or antiinflammatory agent, for treating solid tumours, adult T-cell
CC   leukaemia/lymphoma, fungal infections, and hyperproliferative vascular
CC   diseases such as restenosis and atherosclerosis. The gene encoding the
CC   210 kDa is called sirolimus effector protein (SEP). The present sequence
CC   represents a glutathione-S-transferase-SEP fragment fusion protein used
CC   to study the activity of SEP
XX
SQ   Sequence 1140 AA;

  Query Match             48.9%;  Score +602;  DB 6;  Length 1140;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches  880;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    915 KYASLCGKSGRIALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTIYAYMKNMWKSARKIDAF 974
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 KYASLCGKSGRIALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTIYAYMKNMWKSARKIDAF 305

Qy    975 QHMQHFVQTMQQCAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ 1034
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    306 QHMQHFVQTMQQCAXHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ 365

Qy   1035 YYSAATEHDRSWYKAWHAWAVWNPFRAVLHYKHQNQCARDEKKCKLRHASGANITNATTAATT 1094
          |||||||||||| ||||||||| |||||||||||| |||||  |||||||||||||||||
```

```
Db    366 YYSAATEHDRSWYXAWHAWAVWNFBAVLHYKHQNXARDEKKKLRHASGANITNATTAATT 425
          |||||||||||| ||||||||| |||||||||| |||||| |||||||||||||||||
Qy   1095 AATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL 1154
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    426 AATATTTASTEGSNSESEAESTXNSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL 485

Qy   1155 SRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVG 1214
          ||||||||||||||||||||||||||| ||||||||||||||| |||||||||||||||
Db    486 SRGNNLQDTLRVLTLWFDYGHWPDVNXALVEGVKAIQIDTWLXVIPQLIARIDTPRPLVG 545

Qy   1215 RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKMCEHSNTLVQQAMMVSE 1274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKMCEHSNTLVQQAMMVSE 605

Qy   1275 ELIRVAILWEHMWHRGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY 1334
          |||||||||||||| ||| |||||||||||||| |||||||||||||||||||||||||
Db    606 ELIRVAILWEHMWHEGLEXASRLYFGERNVKGMFXVLEPLHAMMERGPQTLKETSFNQAY 665

Qy   1335 GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVSPKLLMCR 1394
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVSPKLLMCR 725

Qy   1395 DLELAVPGTYDPNOPIIRIQSIAPSLQVITSKQRPKLTLMGSNGHBFVFLLKGHEDLRQ 1454
          |||||||||| || |||||||||||| ||||||||||||||||||| |||||||||||
Db    726 DLELAVPGTYXPNQPIIRIQSIAPSLXVITSKQRPKLTLMGSNGHEFVFLLKGHEDLRQ 785

Qy   1455 DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR 1514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR 845

Qy   1515 EKKKILLNIEHR.MLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEWWF 1574
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    846 EKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEWWF 905

Qy   1575 DRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKLIHPDGDCFBVAMTREKFPEK 1634
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||| ||||
Db    906 DRRTNYTRSLAVMSMVGYILGLGXDRHPSNLMLDRLSGKLIHPDGDCFBVAMTRXKFPEK 965

Qy   1635 IPFRLRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT 1694
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||| |
Db    966 IPFRLRMLTXAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMXT 1025

Qy   1695 NTKGNKRSRTRTDSYSAGQSVEILDGVRLGEPAHKKTGTTVPESIHSFIGDGLVKPEALN 1754
          |||||||||||||||||||||||||| ||| |||||||||||||||||||||||||||||
Db   1026 NTKGNKRSRTRTDSYSAGQSVEILDGVELGXPAHKKTGTTVPESIHSFIGDGLVKPEALN 1085

Qy   1755 KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLKQATSHENLCQCYIGWCPFW 1809
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   1086 KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLKQATSHENLCQCYIGWYPFW 1140
```

```
RESULT 12
AB891303
ID   AB891303 standard; protein; 2467 AA.
XX
AC   AB891303;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 514.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP009892.
XX
PR   28-AUG-2001; 2001WO-EP009892.
XX
```

```
XX
DE   GST-SEP fusion protein fragment.
XX
KW   Glutathione-S-transferase; GST; FK506 binding protein; FKBP; PCR;
KW   rapamycin; FKBP-rapamycin binding protein; Molt4 cell; amplify;
KW   fusion protein; GST-FKBP12; immunomodulatory agent; primer;
KW   antitumour agent; detection; antisense DNA; immune system.
XX
OS   Homo sapiens.
XX
PN   EP676471-A2.
XX
PD   11-OCT-1995.
XX
PF   07-MAR-1995;   95SEP-00301475.
XX
PR   08-MAR-1994;   94US-00207975.
PR   26-SEP-1994;   94US-00312023.
PR   13-FEB-1995;   95US-00384524.
XX
PA   (AMHP ) AMERICAN HOME PROD CORP.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Molnar-Kimber KL, Failli AA, Caggiano TJ, Nakanishi K, Chen Y;
XX
DR   WPI; 1995-346091/45.
DR   N-PSDB; AAT00771.
XX
PT   New effector proteins of rapamycin - which bind to a glutathione-S-
PT   transferase-FK506 binding protein-rapamycin complex.
XX
PS   Example 2; Page 29-35; 44pp; English.
XX
CC   This sequence represents a fragment of a fusion protein between
CC   glutathione S tranferase (GST) and sirolimus effector protein (SEP) gene.
CC   It represents the fragment begining at the linker sequence between the
CC   GST and SEP45 fragment. The cDNA encoding this sequence was isolated from
CC   Molt 4 human T-cell leukaemia cells (ATCC CRL 1582). The sep gene was
CC   amplified in five fragments using the primers given in AAT00756-69. The
CC   amplified SEP gene was inserted into a vector already containing the GST
CC   gene and the fusion protein expressed. The fusion protein was used in the
CC   isolation of a protein of mammalian origin which binds a glutathione-S-
CC   transferase (GST)-FK506 binding protein (FKBP)-rapamycin complex. The
CC   FKBP-rapamycin binding protein and corresponding DNA was isolated from
CC   Molt4 cells using a complex of the fusion protein GST-FKBP12 and
CC   rapamycin. The isolated proteins have molecular weights of 125, 148, 208
CC   and 210 kD. They can be used for identifying an immunomodulatory, or an
CC   antitumour agent. They can also be used in the detection of rapamycin,
CC   rapamycin analogues or metabolites when complexed with FKBP. Antisense
CC   DNA can be used to modulate the immune system of a mammal
XX
SQ   Sequence 1140 AA;

Query Match             49.8%;  Score 4685;  DB 2;  Length 1140;
Best Local Similarity   99.8%;  Pred. No. 0;
Matches  893;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    915 KYASLCGKSGRIALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAF 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 KYASLCGKSGRIALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAF 305

Qy    975 QHMCGHFVQTMQCQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 QHMCGHFVQTMQCQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ 365

Qy   1035 YYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQOARDEKKKLRHASGANITNATTAATT 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 YYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQPRDEKKKLRHASGANITNATTAATT 425

Qy   1095 AATATTTASTEGSNSESRAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 AATATTTASTEGSNSESRAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISL 485
```

```
Qy   1155 SRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVG 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 SRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVG 545

Qy   1215 RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSE 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 RLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCEHSNTLVQQAMMVSE 605

Qy   1275 ELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 ELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY 665

Qy   1335 GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTSLELQYVSPKLLMCR 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTSLELQYVSPKLLMCR 725

Qy   1395 DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDIRQ 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQ 785

Qy   1455 DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYR 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYR 845

Qy   1515 EKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWF 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    846 EKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWF 905

Qy   1575 DRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEK 1634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 DRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEK 965

Qy   1635 IPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT 1694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    966 IPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT 1025

Qy   1695 NTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN 1754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 NTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN 1085

Qy   1755 KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1086 KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWTPFW 1140
```

```
RESULT 11
ABU61980
ID   ABU61980 standard; protein; 1140 AA.
XX
AC   ABU61980;
XX
DT   26-AUG-2003  (first entry)
XX
DE   Human sirolimus effector protein/GST fusion protein.
XX
KW   Human; Rapamycin effector protein; glutathione-S-transferase;
KW   GST-FK-506 binding protein (FKBP)-rapamycin complex; anti-restenosis;
KW   anti-tumour; transplantation rejection; graft versus host disease;
KW   autoimmune disease; lupus; myasthenia gravis; multiple sclerosis;
KW   rheumatoid arthritis; Type I diabetes; inflammatory disease; psoriasis;
KW   dermatitis; eczema; seborrhea; diarrhea; inflammatory bowel disease;
KW   pulmonary inflammation; asthma; eye uveitis; immunosuppressive;
KW   antiinflammatory agent; tumour; adult T-cell leukaemia; lymphoma;
KW   fungal infection; hyperproliferative vascular disease; restenosis;
KW   atherosclerosis; sirolimus effector protein; SEP; GST-SEP;
KW   glutathione-S-transferase.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 1..1140
FT                   /note= "All Xaa reidues have been input by the indexer to
```

```
Query Match          49.8%;  Score 4690;  DB 5;  Length 1140;
Best Local Similarity 99.9%;  Pred. No. 0;
  Matches  894;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    0;

Qy    915 KYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAF  974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 KYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAF  305

Qy    975 QEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ 1034
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 QEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ  365

Qy   1035 YYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATT 1094
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 YYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATT  425

Qy   1095 AATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFPRSISL 1154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 AATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFPRSISL  485

Qy   1155 SRGNNLQDTLRVLLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVG 1214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 SRGNNLQDTLRVLLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVG  545

Qy   1215 RLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKLLKNMCEHSNTLVQQAMMVSE 1274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 RLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKLLKNMCEHSNTLVQQAMMVSE  605

Qy   1275 ELIRVAILWEHMWHEGLEEASRLIYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY 1334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 ELIRVAILWEHMWHEGLEEASRLIYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAY  665

Qy   1335 GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVSPKLLMCR 1394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 GRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVSPKLLMCR  725

Qy   1395 DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVLLKGHEDLRQ 1454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 DLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVLLKGHEDLRQ  785

Qy   1455 DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR 1514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 DERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYR  845

Qy   1515 EKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWF 1574
          |||||||||||||||||:||:||||||||||||||||||||||||||||||||||||||||
Db    846 EKKKILLNIEHRIMLRVAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWF  905

Qy   1575 DRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKLIHIDFGDCFEVAMTREKFPEK 1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 DRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKLIHIDFGDCFEVAMTREKFPEK  965

Qy   1635 IPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT 1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    966 IPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDT 1025

Qy   1695 NTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN 1754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 NTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALN 1085

Qy   1755 KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1086 KKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWYPFW 1140
```

RESULT 10
AAR81731
ID   AAR81731 standard; protein; 1140 AA.
XX
AC   AAR81731;
XX
DT   29-MAY-1996 (first entry)

```
Qy   1720 GVELGEPAHKK---TGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHD 1776
          |||| ||||||   || ||||||||||||||||||  ||| : ||  ||||||||  ||
Db   2384 SVEDSLPMAKSKPYDPTLQQGGLHNNVAD-----ETNSKASQVIKRVKCKLTGTDFQTE 2437

Qy   1777 DTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
          :::  :||||||:|||:||||||||||||||||
Db   2438 KSVNEQSQVELLIQQATNNENLCQCYIGWCPFW 2470
```

RESULT 9
AAE13360
ID   AAE13360 standard; protein; 1140 AA.
XX
AC   AAE13360;
XX
DT   12-FEB-2002 (first entry)
XX
DE   pGEX-Sep45 plasmid protein.
XX
KW   Rapamycin; FK506 binding protein 12; FKBP12; cytostatic;
KW   transplant rejection; autoimmune disease; lupus; rheumatoid arthritis;
KW   diabetes mellitus; myasthenia gravis; multiple sclerosis; vasotropic;
KW   inflammatory disease; psoriasis; dermatitis; eczema; seborrhoea; tumour;
KW   lymphoma; fungal infection; eye uveitis; adult T-cell leukaemia;
KW   restenosis; atherosclerosis; hyperproliferative vascular disease;
KW   antiarthritic; antirheumatic; fungicide; antiarteriosclerotic;
KW   immunosuppressive; neuroprotective; ophthalmological.
XX
OS   Unidentified.
XX
PN   US6313264-B1.
XX
PD   06-NOV-2001.
XX
PF   06-JUN-1995; 95US-00471112.
XX
PR   08-MAR-1994; 94US-00207975.
PR   26-SEP-1994; 94US-00312023.
PR   13-FEB-1995; 95US-00384524.
XX
PA   (AMHP) AMERICAN HOME PROD CORP.
PA   (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI   Caggiano TJ, Chen Y, Failli AA, Molnar-Kimber KL, Nakanishi K;
XX
DR   WPI; 2002-017010/02.
DR   N-PSDB; AAD22154.
XX
PT   New isolated rapamycin effector protein which binds to complex of FK-506
PT   binding protein 12 and rapamycin useful for identification, design and
PT   synthesis of immunomodulatory, anti-restenosis or anti-tumor agents.
XX
PS   Example 2; Col 39-46; 28pp; English.
XX
CC   The invention relates to Rapamycin-FK506 binding protein 12 (FKBP12)
CC   binding proteins of mammalian origin. Rapamycin effector protein is
CC   useful for identification, design and synthesis of immunomodulatory, anti
CC   -restenosis or anti-tumour agents. rapamycin effector protein is useful
CC   in screening assays, such as enzyme inhibitor assays and binding assays,
CC   to identify endogenous complexes and ligands and novel exogenous
CC   compounds (like rapamycin) which modulate their function. The compounds
CC   identified by rapamycin effector protein is useful for treating
CC   transplant rejection, autoimmune diseases such as lupus, rheumatoid
CC   arthritis, diabetes mellitus, myasthenia gravis and multiple sclerosis,
CC   inflammatory diseases such as psoriasis, dermatitis, eczema, seborrhoea,
CC   inflammatory bowel disease and eye uveitis, solid tumours, adult T-cell
CC   leukaemia, lymphoma, fungal infections and hyperproliferative vascular
CC   diseases such as restenosis and atherosclerosis. The present sequence is
CC   a pGEX-Sep45 plasmid protein used in the exemplification of the invention
XX
SQ   Sequence 1140 AA;

```
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC, Adams M, .Li PWD, Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL05350.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 10533; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2470 AA;

Query Match             57.3%;  Score 5389.5;  DB 4;  Length 2470;
Best Local Similarity   58.5%;  Pred. No. 0;
Matches 1073;  Conservative 244;  Mismatches 413;  Indels 103;  Gaps  17;

Qy        1 LEHSCIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA 60
            ::|:: |||||||:::::||| : |:| :||||||:|:|| :||||: : ||| :||
Db      717 LKYSGMSRNKEQSAKMLDHLVISTPRLLISSYMNPILKALVPKLH--EPESNPGVILNVLR 774

Qy       61 TIGHELAQVSG--LEMRKWVDELFFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPY 118
            ||| : |||:: || :|| ||:::||:| |:||| :|: |||| ||||:|| | | ||
Db      775 TIGDLAEVNGGSDEMELWADDLLSILLEMLGDAGSPDKRGVALWTLGQLISATGRVVTPY 834

Qy      119 RKYPTLLEVLLNFLKTEQNQGTRREALRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLS 178
            :|| : ||| ::|||:: | | :|||||||||:||  ||||| || ||| : |
Db      835 HKYPVLIDILINFLKTEQRSIRREFTRVLGLLGAMDPYKHKMNKGLIDSQKDNVLIAYS 894

Qy      179 ESKSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFI 238
             :|:: : |:: |||:|:|  :| :||| | :||||||:  : |:||||| |||||:|
Db      895 DGK-VDESQDISTABELLVNMGN-ALDEYYPAVAIAALMRILRDPTLSTRHTSVVQAVTFI 952

Qy      239 FKSIGLKCVQFLPQYMPTTLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDBIVTLM 298
            |:| : |||| |||| | |||||:||:| |: :| |:||:| |:| ||:|||: :::
Db      953 FQSLGIKCVPVIAQVLPNLLDNVRTADNNLREFLFQQLAIIVAFVKLHIISYMGDIFKLI 1012

Qy      299 REFWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMEDNSPGRIVSIKLLA 358
            ::| :::|:|:|:|||| :|||| | | ||||| || ||||||: |:||| :|| |
Db     1013 KEFWTINTPLQNTLNLIEQIAVALGCEFRDYLAELIPQILRVLQHDNSKDRMVTRRLLQ 1072

Qy      359 AIQLFGANLDDYIHLLLPPIVKLFDAPEAPLPSRKAALETVDRLITESLDFTDYASRIIHP 418
            ||: | | |||| :||| |||||| | :|:| || | ||||: :|||:::|||:|||||||
Db     1073 ALQKFGSTLGYYLPILLPPIVKLFDSPYVPQQVSMVALETINNLACQLDFTDFSSRIIHP 1132

Qy      419 IVRTLDGSPELRSTAMDTLSSIVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVK 478
            |||:|| |  ||::   |||:: |:|| :| |: :| : |:  |: :|:|| |:||||
Db     1133 LVRVLDAEPELRDQAMTTIRSLAKQLGKKYIVFVPMVQRTLNKHRIVDPEYEBLLSKIKS 1192

Qy      479 GYTLADEEEDPLIYCHRMLRSGGQDALASGPVETGPM------KKLHVSTINLQKAWG 530
             :||:| |     ||   |:|| | |  | :  |: |    ::||||:|||| |||
Db     1193 CSTLADS-------YGAGESELRPSRFKNNEPFVTDRNSNNKNLQVTTNELRTAWQ 1241

Qy      531 AARRVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALLAQAYNPMARDLFNAAFVSCWSELN 590
             |||::|:|| : |:| : |:|||||::|:||:| | ||||| ||| : |:: |:: || :
Db     1242 VTRRVSKDDWVEWLKRLSIGLLKEPSHALRACRSLAQEYDTLLRDLFNAAFISCWTELS 1301

Qy      591 EDCQQDELIRSIELALTSQDIAEVTQTLLNLAEPMEHSDKGPLPLRDDNGIVLLGERAAKC 650
            |: | ||| :: |:|:| |: |:|  ::|: ||   |||  ||| || |  :||| |
Db     1302 PDLKNELTQSLIQALQVTDMPEITQTIINLAEFMEHCDRDPIPIETK----LLGTRAMAC 1357

Qy      651 RAYAKALHYKELFFQKGPTPAILELSLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWY 710
            ::|||||:|| || | || :| ||::::||||||:| |:|  |:|: |: || || |:| |
Db     1358 RAYAKALRYKEEFLLREDSQVFESLILINNKLQQREAAEGLLTRYRNAAELNVQGRWY 1417

Qy      711 EKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQA 770
            ||||:|| ||: |::|:  |||  |:||| ::|| |:||: | |::  |  | :: :::|:
Db     1418 EKLHNWDEALEHYERNLKTDSSDLEARLGHMRCLEALGDWSELSNVTKHEWENFGTRAKS 1477

Qy      771 KMARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDL 830
            :   | |||| || ||: |  ||  ||:|  | | :|||||||:  |:  || | |: ||
Db     1478 RAGPLAAVAAWGLQDWEAMREYVRCIPEDTQDGSYYRAVLAVHHDDFETAQRLIDETRDL 1537

Qy      831 LDAELTAMAGESYRAYGAMVSCHMLSELEVIQYKLVPEREBIRQIWWERLQGCCRIV 890
            ||  |::||| ||||||||:| ||:||||:||||:| | |: |  ::||:||:|| :|
Db     1538 LDTELTSNAGESYERAYGAMVCVQMLAELEEVIQYKLIPERREPLKTMWWKRLQGGQRLV 1597

Qy      891 EDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLP 950
            ||| | :|::| |::| ||:|||||||||| ||  || ||:|| || |||||||:| :
Db     1598 EDWRRIIQVHSLVVKPHEDIHTWLKYASLCRKSGGLLH:SHKTLVMLLGTDPKLNPNQPLP 1657

Qy      951 TVHPQVTYAYMKNMWKSARKIDAPQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLWARC 1010
            |:: ||| |||| ||:   :::|:| |::: |:   :| |::   | : : |: :||: |
Db     1658 CNQPQVTYAYTKYMAANNLQRAVEQLTHFVSTYSQELS---CLPPEALKQQDQRLMARC 1714

Qy     1011 FLKLGEWQLNLQ-GINESTIPKVLQYYSAATEHDRSWTYKAWHAWAVMNFEAVLHYKHQNQ 1069
             |||:||||||| | ||| |||||| |:|| ||  ||| ||| | :  :  ||:| | ||
Db     1715 YLRMATWQNKLQDSIRPDAIQGALECFEKATSYDPNWTKAWHLWAYMNFKVV---QAQKS 1771

Qy     1070 ARDEKKKLRHASGANITNATTAATTAASTEGSNSBSEAESTENSPTPSPLQKKV 1129
            | :  |                   : |::| ::
Db     1772 ALDKQQ----PPGASM---------------GMTMGGGLDSDL-- 1795

Qy     1130 TEDLSKTLLMYTVPAVQGFFRSILSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKA 1189
             :  |  |: |||| |||||||:|:|| ||||||:||:|||:|||: :||| :| |:|
Db     1796 -----MIIQRYAVPAVQGFFRSISLIKGNSLQDTLRLLTLWFDYGNHAEVVEALLSGMKL 1850

Qy     1190 IQIDTWLQVIPQLIARIDTPRPIVGRLHQLLTDIGRVHPQALIYPLTVASKSTTARHN 1249
            |: ||||:|||||| ||| || :| |||:| |||||:||| :|||||:|||| |: |::
Db     1851 IEINTWLQVIPQLIARIDTHRQLVGQLIHQLLMDIGKNHPQALVVPLTVASKSASLARRN 1910

Qy     1250 AANKILKNMCESNTLVQQAMMVSEELIRVALIWHEMWHBGLEEASRLYFGERNVKGMFE 1309
            ||  |||:: |:|||:| |:|:|||||| |::|| | |:||||||||:|| | ||:||:
Db     1911 AAFKILDSMRKHSPTLVEQAVMCSEELIRVAIIWHEQWHEGLEEASRLYFGDRNVKGMFE 1970

Qy     1310 VLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFR 1369
            :|: |||||||||||||||||:|| |:| | |||:|:|| |||: |: |:| |:|||:
Db     1971 ILEPLHAMLERGPQTLKETSFSQAYGRELTEAYEWSQRYKTSAVVMDLDRAWDIYYHVFQ 2030

Qy     1370 RISKQLPQLTSLELQYYSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRP 1429
            :||:||| |||||||:|||:| ||||||||||||||||| ||||:||| ||||::||||
Db     2031 KISRQLPQLTSLELPYVSPKLMTCKDLELAVPGSYNPGQELIRISIIKTNLQVITSKQRP 2090

Qy     1430 RKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLPFGLVNTLLANDPTSLRKNLSIQRYAVIP 1489
            ||| | || || ||| |||||||||||||||| |||:|| |:|||:|:||: | |||||||
Db     2091 RKLCIRGSNGKDYMYLLKGHEDLRQDERVWQLFSLVNTLLLDDPTFRRNLAIQRYAVIP 2150

Qy     1490 LSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRWAPDYDHLTLMQKVEVFEH 1549
            |||||||||||||||||||||||||:||| |||||:| |||::| |||| ||||:|||| |
Db     2151 LSTNSGLIGWVPHCDTLHTLIRDYRDKKKKVPLNQBHRTWLNFAPDYDHLTLMQKVEVFEH 2210

Qy     1550 AVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRL 1609
            |  |:  || |||:|||||| | |::|| |||||||||||||||:|:|||||||||| |
Db     2211 ALGCTQSDDIAKLIWLKSPSSELWFERRNNYTRSLAVMSMVGYILGLGDRHPSNLMLDRM 2270

Qy     1610 SGKILEIHIDFGDCFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLRE 1669
            |||||:||||||||||||||:|||||||| || |:|| ||: |:|:|| |:||:|: ::||
Db     2271 SGKILHIDFGDCFEVAMTREKFPEKIPFRITRMLIKAMEVTGIFGTYRRTCESVMLVLRR 2330

Qy     1670 HKDSVMAVLEAFVVDPLLNWRLMDTNTKGNKGSRTRTDSYS-----------AGQSVEILD 1719
            |:||:|: | ||:|||||||||:|      ::||::
Db     2331 NKDSLMAVLEAFVVDPLLNWRLLDVDKKGN------DAVAGAGAPGGRGGSGMQDSLSN 2383
```

```
Db   741 LEHSGIGRIKEXSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA 800

Qy    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVVEPYRK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVVEPYRK 860

Qy   121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKCHKVNIGMIDQSRDASAVSLSES 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKCHKVNIGMIDQSRDASAVSLSES 920

Qy   181 KSSQDSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 240
         ||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   921 KSSQDSXYSTSEMLVNMGNLPLXEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 980

Qy   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
         |||||||||| |||||||||||||||||||||||||||||||||||||||||||||| |
Db   981 SLGLKCVQFLPXVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRX 1040

Qy   301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPEMLRVFMHDNSPGRIVSIKLLAAI 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1041 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPEMLRVFMHDNSPGRIVSIKLLAAI 1100

Qy   361 QLEGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1101 QLEXANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQTFIPMVNKVLVHRINHQRYDVLICRIVKGY 480
         ||||||| |||||||||||||||||||||| |||||||||||||||||||||||||||
Db  1161 RTLDQSPXLLRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVHRINHQRYDVLICRIVKGY 1220

Qy   481 TLADEEEDPLIYQHRMLRSGQGDALASGWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 540
         ||||||||||||||||||||||||||| 
Db  1221 TLADEEEDPLIYQHRMLRSGQGDALASXPVETGPMKKLHVSTIXLQXAWGAARRVSKDDW 1280

Qy   541 LEWLRRISLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
         ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  1281 LEWLRRISLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSXINEDQQDELIRS 1340

Qy   601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db  1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRXDNGIVLLGERAAKCRAYAKALYK 1400

Qy   661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
         ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  1401 ELEFQKGPTPAILESLISINNKLQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy   721 VAYDKKMDTNKDDPELMGRMRCLEALGEWGQLHQQCCEKKWTLVNDETQARVARMAAAAA 780
         ||||||||||| |||||||||||||| |||||||||||||||||||||| |||||||||
Db  1461 VAYDKKMDTNKXDPELMGRMRCLEALXEWGQLHQQCCEKKWTLVNDETQAKVARMAAAAA 1520

Qy   781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 900
         |||||||||||||||||| |||||||||||||||||||||||||||||||  |||||||
Db  1581 ESYSRAYGAMVSCHMLSELXEVIQYKLVPERREIIRQIWERLQGCQRIVXXWQKILMVR 1640

Qy   901 SLVVSGPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 960
         ||||| |  ||||||||||||||| ||||| |||| |||| |||||||||||||||||||
Db  1641 SLVVSPHXDMRTWLKYASLCGKSXRLALAAHKTLVILLGVXPSRQLDHPLPTVHPQVTYAY 1700

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHQELHKLMARCFLKLGGEWQLN 1020
         |||||||||| |||||||||||||||||||||||||||| |||||||||||||||||||
Db  1701 MKNMWKSARKIXAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGGEWQLX 1760

Qy  1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVNNFEAVLEYKHQNQARDEKKKLRHA 1080
         |||||||||||||||||||||||||||||||||||||||||||| |||| |||||||||
Db  1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVNNFEAVLHYKHQXQARDEKKKLRHA 1820

Qy  1081 SGANITNATTAATTAATATTTAASTEGSNSESEABSTENSPTSPLQKKVTEDLSKTLLMY 1140
         ||| ||||||||||| ||||||||||||||||||||||||||||||||||| |||||||
Db  1821 SGAXITNATTAATTAAATATTTAASTEGSNSESEABSTENSPTSPLQKKVTEXLSKTLLMY 1880

Qy  1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQDTWLQVIP 1200
         ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVKEALVEGVKAIQDTWLQVIP 1940

Qy  1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKQMCE 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKQMCX 2000

Qy  1261 HSNTLVQQAMMVSEELIRVALLMHEMWHEGLEEASRLYFGERNVKGMPEVLEPLHAMMER 1320
         ||||||||||||||||||||| |||||||||| |||||||||||||||||||||||||||
Db  2001 HSNTLVQQAMMVSEELIRVAIIMHEMWHEGLXEASRLYFGERNVKGMPEVLEPLHAMMER 2060

Qy  1321 GPQTLKETSFNQAYGRDIMEAQEWCRKYMKSGNVKDLTQAWDLIYHVFRRISKQLPQLTS 1380
         ||||||||||||||||| |  ||||||||||||| |||||||||||||||||||||||||
Db  2061 GPQTLKETSFNQAYGRDLMXAQEWCRKYMKSGNVKXLTQAWDLIYHVFRRISKQLPQLTS 2120

Qy  1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMXSNGH 2180

Qy  1441 EFVFLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2181 EFVFLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy  1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
         ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFXHAVNNTAGDDLA 2300

Qy  1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy  1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy  1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTVPESIH 1740
         ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELXEPAHKKTGTVPESIH 2480

Qy  1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRXFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy  1801 CYIGWCPFW 1809
         ||| | |||
Db  2541 CYIXWYPFW 2549
```

```
RESULT 8
ABB61247
ID   ABB61247 standard; protein; 2470 AA.
XX
AC   ABB61247;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 10533.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
```

Db   1761 LQGINRSTIPKVLQYYSAATEHDRSWYKAMHRAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081 SGANINATTAAATTAATTAASTEGSNSESFAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
             ||||||||||||||| ||||||| :|
Db   1821 SGANINATTAATTAATTAASAAAATSTEGSNSESFAESNESSPTPSPLQKKVTEDLSKTLLIY 1880

Qy   1141 TVPAVOGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881 TVPAVOGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy   1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKIIKNMCE 2000

Qy   1261 HSNTLVQQAMMVSEELIRVAILWEHMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2001 HSNTLVQQAMMVSEELIRVAILWEHMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy   1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy   1381 LELQYVSPKLLMCRDLELAVPGTTYDPNCQPIIRIQSIAPSLQVITSKQRPRKLJTLMGSNGH 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2121 LELQYVSPKLLMCRDLELAVPGTTYDPNQTIIRIQSIAPSLQVITSKQRPRKLJTLMGSNGH 2180

Qy   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRTCHVMEVLREHKDSVMAVLEA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDRNYRTCHVMEVLREHKDSVMAVLEA 2420

Qy   1681 FVYDPLLNWRLMDTNYTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2421 FVYDPLLNWRLMDTNAKGNKRSRTRTDSYSAGQSVEILDGVELGBPAHKKTGTTVPESIH 2480

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVFTQVELLIKQATSHENLCQ 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVFTQVELLIKQATSHENLCQ 2540

Qy   1801 CYIGWCPFW 1809
             |||||||||
Db   2541 CYIGWCPFW 2549

RESULT 7
ABU61979
ID   ABU61979 standard; protein; 2549 AA.
XX
AC   ABU61979;
XX
DT   26-AUG-2003 (first entry)
XX
DE   Human sirolimus effector protein.
XX
KW   Human; Rapamycin effector protein; glutathione-S-transferase;
KW   GST-FK-506 binding protein (FKBP)-rapamycin complex; anti-restenosis;
KW   anti-tumour; transplantation rejection; graft versus host disease;
KW   autoimmune disease; lupus; myasthenia gravis; multiple sclerosis;
KW   rheumatoid arthritis; type I diabetes; inflammatory bowel disease; psoriasis;
KW   dermatitis; eczema; seborrhea; inflammatory disease; psoriasis;
KW   pulmonary inflammation; asthma; eye uveitis; immunosuppressive;
KW   antiinflammatory agent; tumour; adult T-cell leukaemia; lymphoma;
KW   fungal infection, hyperproliferative vascular disease; restenosis;
KW   atherosclerosis; sirolimus effector protein; SEP.
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   misc_difference 1 .2549
FT   /note "All Xaa residues have been input by the indexer
FT   to restore the number of residues to that stated in the
FT   specification (residues omitted by an apparent printing
FT   error)"
XX
PN   US2003032775-A1.
XX
PD   13-FEB-2003.
XX
PF   13-SEP-2001; 2001US-00950634.
XX
PR   08-MAR-1994;   94US-00207975.
PR   26-SEP-1994;   94US-00312023.
PR   13-FEB-1995;   95US-00384524.
PR   06-JUN-1995;   95US-00471112.
XX
PA   (AMHP) AMERICAN HOME PROD CORP.
XX
PI   Caggiano TU,   Chen Y,   Failli AA,   Molnar-Kimber KL,   Nakanishi K;
XX
DR   WPI; 2003-492083/46.
DR   N-PSDB; ACA63012.
XX
PT   Novel effector protein of rapamycin of mammalian origin for identifying
PT   immunomodulatory or anti-tumor agent, binds glutathione-S-transferase-FK-
PT   506 binding protein-rapamycin complex.
XX
PS   Disclosure; Page 15-21; 32pp; English.
XX
CC   The invention relates to an effector protein of rapamycin of mammalian
CC   origin having a molecular weight of about 125, 148, 208 or 210 kDa which
CC   binds to a glutathione-S-transferase (GST)-FK-506 binding protein (FKBP)-
CC   rapamycin complex. Also included are a recombinantly produced protein
CC   comprising a partial amino acid sequence of one of the proteins listed
CC   above, cDNA/DNA of mammalian origin which encodes the proteins and
CC   antisense DNA or RNA derived from cDNA/DNA clone. The proteins are useful
CC   for identifying an immunomodulatory or anti-tumour agent, for detecting,
CC   in a biological sample, rapamycin, rapamycin analogues or metabolites.
CC   Antisense RNA or DNA is useful for modulating the immune system of a
CC   mammal in need of it. The proteins are useful for identifying, designing
CC   or synthesising immunomodulatory, anti-restenosis or anti-tumour agents,
CC   in screening assays, such as enzyme inhibitor assays and binding assays,
CC   to identify endogenous complexes and ligands and novel exogenous
CC   compounds (e.g. rapamycin) which modulate their functions, and in assays
CC   to identify a compound with therapeutic benefit for restenosis
CC   immunomodulation and as antitumour agents. Antisense DNA or RNA is useful
CC   for treating transplantation rejection, graft versus host disease,
CC   autoimmune diseases such as lupus, myasthenia gravis, multiple sclerosis,
CC   rheumatoid arthritis, type I diabetes, and diseases of inflammation such
CC   as psoriasis, dermatitis, eczema, seborrhea, inflammatory bowel disease,
CC   pulmonary inflammation, asthma, and eye uveitis, as an immunosuppressive
CC   or antiinflammatory agent, for treating solid tumours, adult T-cell
CC   leukaemia/lymphoma, fungal infections, and hyperproliferative vascular
CC   diseases such as restenosis and atherosclerosis. The gene encoding the
CC   210 kDa is called sirolimus effector protein (SEP). The present sequence
CC   is the SEP protein
XX
SQ   Sequence 2549 AA;

Query Match   97.5%;   Score 9176;   DB 6;   Length 2549;
Best Local Similarity   98.1%;   Pred. No. 0;
Matches 1774;   Conservative   0;   Mismatches   35;   Indels   0;   Gaps   0;

Qy   1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDPNPGVINNVLA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Query Match   99.1%;   Score 9325;   DB 6;   Length 2549;
Best Local Similarity   99.0%;   Pred. No. 0;
Matches 1791;   Conservative   5;   Mismatches   13;   Indels   0;   Gaps   0;

```
Qy      1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPNPGVINNVLA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPNPGVINNVLA 800

Qy     61 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTIGQLVASTGYVVEPYRK 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801 TIGELAQVSGLEMRKWVDELFVIIMDMLQDSSLLAKRQVALWTIGQLVASTGYVVEPYRK 860

Qy    121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    861 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy    181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIPRDQSLSHHTMVVQAITPIFK 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIPRDQSLSHHTMVVQAITPIFK 980

Qy    241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy    301 FWVMNTSIQSTILLIEQIVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 FWVMNTSIQSTILLIEQIVALGGEFKLYLPQLIPHMLRVFMDNSQGRIVSIKLLAAI 1100

Qy    361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLLLPPIVKLFDAPEVPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy    421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy    481 TLADEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 TLADEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy    541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy    601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy    661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401 ELEFQKGPTPAILESLISINNKLQQPEAASGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy    721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy    781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy    841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ESYSRAYGAMVSCHMLSELEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 1640

Qy    901 SLVVSPHEDMRIWLKYASLCGKSSRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 SLVVSPHEDMRIWLKYASLCGKSSRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy    961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKIGEWQLN 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKIGEWQLN 1760

Qy   1021 LQGINEST-PKVLQYYSAATEHDRSWYKAWHAWAVMNPEAVLHYKHQNQARDEKKKLRHA 1080
```

```
Qy   1561 KLIWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2301 KLIWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy   1681 FVTDPLLNWRLMDTNTKGNKRSRTRIDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2421 FVTDPLLNWRLMDTNAKGNKRSRTRIDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy   1801 CYIGWCPFW 1809
            |||||||||
Db   2541 CYIGWCPFW 2549
```

RESULT 6
ABU08621
ID   ABU08621 standard; protein; 2549 AA.
XX
AC   ABU08621;
XX
DT   27-MAY-2003   (first entry)
XX
DE   Rat (rapamycin and FKBP target 1) RAFT1.
XX
KW   RAFT1; rapamycin and FKBP12 target 1; rat.
XX
OS   Rattus rattus.
XX
PN   US6492106-B1.
XX
PD   10-DEC-2002.
XX
PF   14-SEP-1994;   94US-00305790.
XX
PR   27-JUN-1994;   94US-00265967.
XX
PA   (UYJO) UNIV JOHNS HOPKINS.
XX
PI   Sabatini DM, Brdjument-Bromage H, Lui M, Tempst P, Snyder SH;
XX
DR   WPI; 2003-340404/32.
DR   N-PSDB; ABX93443.
XX
PT   Novel isolated, purified cDNA molecule which encodes a rapamycin and
PT   FKBP12 target, referred as RAFT1 protein, useful as probe or primer for
PT   identifying other mammalian RAFT proteins.
XX
PS   Claim 1; Col 25-38; 46pp; English.
XX
CC   The invention describes an isolated, purified cDNA molecule (I) which
CC   encodes RAFT1 (rapamycin and FKBP12 target 1). An isolated intron free
CC   DNA is useful as a probe for isolating a DNA molecule encoding a
CC   mammalian RAFT protein, which involves probing a rat cDNA sequences with
CC   a probe which comprises at least 15 contiguous nucleotides selected from
CC   (SI), and isolating a rat cDNA molecule which hybridises to the probe,
CC   contains a complete open reading frame encoding a polypeptide of 2550
CC   amino acids, and encodes a rat RAFT protein, where the rat RAFT protein
CC   binds to FKBP12 in the presence of 1-10 nM rapamycin but not in the
CC   absence of 1-10 nM rapamycin. (I) is useful for generating probes which
CC   are used to screen library of mammalian DNA molecules. (I) is also useful
CC   as a primer. This is the amino acid sequence of rat RAFT1 (rapamycin and
CC   FKBP target 1)
XX
SQ   Sequence 2549 AA;

XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P42346.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
CC   injury (CCI) and spared nerve injury (Chung), chronic constriction
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2549 AA;

Query Match   99.2%; Score 9341; DB 7; Length 2549;
Best Local Similarity   99.1%; Pred. No. 0;
Matches 1793; Conservative   5; Mismatches   11; Indels   0; Gaps   0;

Qy   1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPLIKALILKDPDPDNPGVINNVLA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPLIKALILKDPDPDNPGVINNVLA 800

Qy   61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGGLVASTGVVVEPYRK 120
              ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db   801 TIGELAQVSGLEMRKWVDELFVIIMDMLQDSSLLAKRQVALWTLGGLVASTGVVVEFYRK 860

Qy   121 YPTLLEVLLNFLKTEQNGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861 YPTLLEVLLNFLKTEQNGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy   181 KSSCQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITPIFK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921 KSSCQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITPIFK 980

Qy   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy   301 FWTMNTSIQSTILLIEQIVVALGGBFKLYLPQLIPEHMLRVFMHDNSPGRIVSIKLLAAI 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 FWTMNTSIQSTILLIEQIVVALGGBFKLYLPQLIPEHMLRVFMHDNSQCRIVSIKLLAAI 1100

Qy   361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLLLPPIVKLFDAPEVPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy   481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 TLADEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy   601 IELALTSQDIAEVTQTLLNLABFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 IELALTSQDIAEVTQTLLNLABFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy   661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db   1401 ELEFQKGPTPAILESLISINNKLQQPEAASGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy   721 VAYDKKMDTNKDDPELMLGMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 VAYDKKMDTNKDDPELMLGMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy   781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLLAHQDLFSLAQQCIDKARDLLDABLTAMAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDABLTAMAG 1580

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKIIMVR 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKIIMVR 1640

Qy   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCPLKGLGEWQLN 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy   1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy   1081 SGANITNATTAATTAATATTTASTEGSNESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
              ||||||||| |||||||||||||||||||||:||||::|:|||||||||||||:|||||:
Db   1821 SGANITNATTTATTASAAAATSTEGSNSESEAESNESSPTPSPLQKKVTEDLSKTLLLY 1880

Qy   1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLMPDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLMPDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy   1201 QLIARIDTPRPIVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKNMCE 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 QLIARIDTPRPIVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKNMCE 2000

Qy   1261 HSNTLVQQAMMVSEBLIRVALLWHEMWHEGLEEAASRLYFGERNVKGMFEVLEPLHAMMER 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2001 HSNTLVQQAMMVSEBLIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy   1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy   1381 LELQYVSPKLLMCRDLEILAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKJTLMGSNGH 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2121 LELQYVSPKLLMCRDLEILAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKJTLMGSNGH 2180

Qy   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy   1501 PHCDTLHLIRDYREKKKILLNIEHRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDLA 1560
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2241 PHCDTLHLIRDYREKKKILLNIEHRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDLA 2300

```
Qy    121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    861 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy    181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy    241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy    301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 1100

Qy    361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLLLPPIVKLFDAPEALLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy    421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIPIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIPIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy    481 TLADEERDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 TLADEERDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy    541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 LEWLRRLTLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy    601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy    661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy    721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQARMARMAAAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQARMARMAAAAA 1520

Qy    781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy    841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 1640

Qy    901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy    961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKILGEWQLN 1020
          ||||||||| ||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1701 MKNMWKSAPKIDAFQHMQHFVQTMQQQAQHPIATEDQQHKQELHKLMARCFLKILGEWQLN 1760

Qy   1021 LQGINESTIPKVLQYYSAAATEHDRSWYKAWHAWAVMNFERAVLHYKHONQARDEKKKLRHA 1080
          ||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
Db   1761 LQGINESTIPKVLQYYSAATEHDRSSYKAWHAWAVMNFERAVLHYKHONQARDEKKKLRHA 1820

Qy   1081 SGANTTNATTNATTAATTAATATTTASTEGSNSESEAESTENSPTSPLQKKVTEDLSKTLLMY 1140
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db   1821 SGANTTNATTNATTAATATTTASTEGSNSESEAESTENSPTSPLQKKVTEDLSKTLLMY 1880

Qy   1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
          |||||||||||||||||||||||||||| |||||||||||||| |||||||||||||||
Db   1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLCFDYGHWPDVNEALEEGVKAIQIDTWLQVIP 1940

Qy   1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 2000

Qy   1261 HSNTLVQQAMMVSEELIRVAIIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
          ||||||| ||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   2001 HSNTLVQDAMMVSEELIRVAIIWHEMWHEGLEEASRLYFGERNVKGMFEVLBPLHAMMER 2060

Qy   1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPFQLTS 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPFQLTS 2120

Qy   1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy   1441 EFVFLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2181 EFVFLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVSVFEHAVNNTAGDDLA 1560
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy   1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy   1681 FVYDPLLNWRLMDINTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy   1801 CYIGWCPFW 1809
          |||||||||
Db   2541 CYIGWCPFW 2549
```

RESULT 5
ADE62170
ID   ADE62170 standard; protein; 2549 AA.
XX
AC   ADE62170;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Rat Protein P42346, SEQ ID NO 8099.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
XX
PN   WO200301675-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO) GEN HOSPITAL CORP.
PA   (FARB) BAYER AG.
XX
PI   Woolf C, D'urso D, Befort K, Costigan M;

```
Qy   1801 CYIGWCQLPFW 1809
          ||||||| ||||
Db   2541 CYIGWYPFW 2549


RESULT 4
AAW56027
ID   AAW56027 standard; protein; 2549 AA.
XX
AC   AAW56027;
XX
DT   28-AUG-1998  (first entry)
XX
DE   FRAP (tor1) protein.
XX
KW   FRAP; tor1 protein; macrolide binding protein; graft versus host disease;
KW   stem cell; gene therapy.
XX
OS   Mammalia.
XX
PN   WO9808956-A2.
XX
PD   05-MAR-1998.
XX
PF   27-AUG-1997;  97WO-US015153.
XX
PR   27-AUG-1996;  96US-0024666P.
XX
PA   (HARD ) HARVARD COLLEGE.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI   Schreiber SL,  Belshaw PJ,  Crabtree G;
XX
DR   WPI; 1998-179441/16.
DR   N-PSDB; AAV28518.
XX
PT   Selective inhibition of proliferation of haematopoietic cells - using
PT   macrolide binding proteins and analogues, useful for treatment of graft
PT   versus host disease.
XX
PS   Disclosure; Page 64-73; 85pp; English.
XX
CC   This polypeptide comprises FRAP (tor1) protein. A claimed method for
CC   selective inhibition of proliferation of a haematopoietic cell comprises
CC   contacting a haematopoietic cell which ectopically expresses a gene
CC   encoding a mutated macrolide binding protein (MBP), which has altered
CC   macrolide-binding specificity relative to the wild-type form MBP, with a
CC   macrolide which selectively induces macrolide dependent inhibition of
CC   proliferation of cells expressing the mutated MBP compared to cells
CC   expressing a wild-type form of the MBP. The MBP is selected from FRAP, an
CC   FK506 binding protein (see AAV28517), a cyclophilin and a calcineurin
CC   (see AAW56028). Also claimed are: an expression construct encoding a
CC   mutated MBP chosen from FRAP, FKBP, cyclophilin and calcineurin; a method
CC   for selectively inhibiting proliferation of a transplanted haematopoietic
CC   cell; a method for treating graft-versus-host disease by selectively
CC   inhibiting proliferation of transplanted haematopoietic cells; and a
CC   method of promoting engraftment and haematopoietic activity of a
CC   haematopoietic stem cell from a donor
XX
SQ   Sequence 2549 AA;

Query Match             99.3%;  Score 9350;  DB 2;  Length 2549;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 1801;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKDPDDPDNPGVINNVLA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKDPDDPDNPGVINNVLA 800

Qy     61 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 860
```

```
Qy    721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQACMARMARMAAAAA 1520

Qy    781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy    841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWERLQGCCRIVEDWQKILMVR 1640

Qy    901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAY 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy    961 MKNMWKSARKIDAFQEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 MKNMWKSARKIDAFQEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy   1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy   1081 SGANITNATTAATTAATTAATTAATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1821 SGANITNATTAATTAATTAATTAATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880

Qy   1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy   1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 2000

Qy   1261 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2001 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy   1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 GPHTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy   1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTIMGSNGH 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTIMGSNGH 2180

Qy   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTMQKEVFEHAVNNTAGDDLA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTMQKEVFEHAVNNTAGDDLA 2300

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy   1621 CFEVAMTREKPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2361 CFEVAMTREKPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy   1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540
```

```
PS   Example 2; Page 16-29; 44pp; English.
XX
CC   This sequence represents the sirolimus effector protein (SEP). The cDNA
CC   encoding this sequence was isolated from Molt 4 human T-cell leukaemia
CC   cells (ATCC CRL 1582) and used in the production of a fusion protein
CC   between glutathione S tranferase (GST) and SEP. The sep gene was
CC   amplified in five fragments using the primers given in AAT00756-69. The
CC   amplified SEP gene was inserted into a vector already containing the GST
CC   gene and the fusion protein expressed (see also AAR81731). The fusion
CC   protein was used in the isolation of a protein of mammalian origin which
CC   binds a glutathione-S-transferase (GST)-FK506 binding protein (FKBP)-
CC   rapamycin complex. The FKBP-rapamycin binding protein and corresponding
CC   DNA was isolated from Molt4 cells using a complex of te fusion protein
CC   GST-FKBP12 and rapamycin. The isolated proteins have molecular weights of
CC   125, 148, 208 and 210 kD. They can be used for identifying an
CC   immunomodulatory, or an antitumour agent. They can also be used in the
CC   detection of rapamycin, rapamycin analogues or metabolites when complexed
CC   with FKBP. Antisense DNA can be used to modulate the immune system of a
CC   mammal
XX
SQ   Sequence 2549 AA;

Query Match             99.8%;  Score 9394;  DB 2;  Length 2549;
Best Local Similarity   99.8%;  Pred. No. 0;
Matches 1806;  Conservative 1;  Mismatches 1;  Indels 2;  Gaps 0;

Qy    1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPPDPNPGVINNVLA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPPDPNPGVINNVLA 800

Qy   61 TIGELAQVSGLEMRKWVDELFTIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  801 TIGELAQVSGLEMRKWVDELFTIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 860

Qy  121 YPTTLEVLLNFLKTEQNGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  861 YPTTLEVLLNFLKTEQNGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy  181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy  241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db  981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRD 1040

Qy  301 FWVMNTSIQSTILLIEQIVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1041 FWVMNTSIQSTILLIEQIVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 1100

Qy  361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1101 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy  421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy  481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGMKKLHVSTINLQKAWGAARRVSKDDW 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1221 TLADEEDPLIYQHRMLRSGQGDALASGPVETGMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy  541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy  601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy  661 ELEFQKGPTPAILESLISNNKLQQPEAAAGVLEYAMKHFGELEFIQATWYEKLHEWEDAL 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1401 ELEFQKGPTPAILESLISNNKLQQPEAAAGVLEYAMKHFGELEFIQATWYEKLHEWEDAL 1460

Db 2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVPRRISKQLPQLTS 2120

Qy 1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy 1441 EPVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2181 EPVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy 1501 PHCDTLHALIRDYREKKKILLNIEIHRIMLRMAPDYDHLTMQKVEVFEHAVNNTAGDDLA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2241 PHCDTLHALIRDYREKKKILLNIEIHRIMLRMAPDYDHLTMQKVEVFEHAVNNTAGDDLA 2300

Qy 1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy 1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy 1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy 1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy 1801 CYIGWCPFW 1809
          ||||| |||
Db 2541 CYIGWYPFW 2549
```

RESULT 3
AAR81730

```
ID   AAR81730 standard; protein; 2549 AA.
XX
AC   AAR81730;
XX
DT   29-MAY-1996  (first entry)
XX
DE   Sirolimus effector protein.
XX
KW   Glutathione-S-transferase; GST; FK506 binding protein; FKBP; PCR;
KW   rapamycin; FKBP-rapamycin binding protein; Molt4 cell; amplify;
KW   fusion protein; GST-FKBP12; immunomodulatory agent; primer;
KW   antitumour agent; detection; antisense DNA; immune system.
XX
OS   Homo sapiens.
XX
PN   EP676471-A2.
XX
PD   11-OCT-1995.
XX
PF   07-MAR-1995;  95EP-00301475.
XX
PR   08-MAR-1994;  94US-00207975.
PR   26-SEP-1994;  94US-00312023.
PR   13-FEB-1995;  95US-00384524.
XX
PA   (AMHP ) AMERICAN HOME PROD CORP.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Molnar-Kimber KL, Failli AA, Caggiano TJ, Nakanishi K, Chen Y;
XX
DR   WPI; 1995-346091/45.
DR   N-PSDB; AAT00770.
XX
PT   New effector proteins of rapamycin - which bind to a glutathione-S-
PT   transferase-FK506 binding protein-rapamycin complex.
XX
```

KW inflammatory disease; psoriasis; dermatitis; eczema; seborrhoea; tumour;
KW inflammatory bowel disease; eye uveitis; adult T-cell leukaemia;
KW lymphoma; fungal infection; hyperproliferative vascular disease;
KW restenosis; atherosclerosis; fungicide; antiarteriosclerotic;
KW antiarthritic; antirheumatic; antiinflammatory; dermatological;
KW immunosuppressive; neuroprotective; ophthalmological.
XX
OS Unidentified.
XX
PN US6313264-B1.
XX
PD 06-NOV-2001.
XX
PF 06-JUN-1995; 95US-00471112.
XX
PR 08-MAR-1994; 94US-00207975.
PR 26-SEP-1994; 94US-00312023.
PR 13-FEB-1995; 95US-00384524.
XX
PA (AMHP) AMERICAN HOME PROD CORP.
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Caggiano TJ, Chen Y, Failli AA, Molnar-Kimber KL, Nakanishi K;
XX
DR WPI; 2002-017010/02.
DR N-PSDB; AAD22153.
XX
PT New isolated rapamycin effector protein which binds to complex of FK-506
PT binding protein 12 and rapamycin, useful for identification, design and
PT synthesis of immunomodulatory, anti-restenosis or anti-tumor agents.
XX
PS Example 2; Col 27-40; 28pp; English.
XX
CC The invention relates to Rapamycin-FK506 binding protein 12 (FKBP12)
CC binding proteins of mammalian origin. Rapamycin effector protein is
CC useful for identification, design and synthesis of immunomodulatory, anti
CC -restenosis or anti-tumour agents. Rapamycin effector protein is useful
CC in screening assays, such as enzyme inhibitor assays and binding assays,
CC to identify endogenous complexes and ligands and novel exogenous
CC compounds (like rapamycin) which modulate their function. The compounds
CC identified by rapamycin effector protein is useful for treating
CC transplant rejection, autoimmune diseases such as lupus, rheumatoid
CC arthritis, diabetes mellitus, myasthenia gravis and multiple sclerosis,
CC inflammatory diseases such as psoriasis, dermatitis, eczema, seborrhoea,
CC inflammatory bowel disease and eye uveitis, solid tumours, adult T-cell
CC leukaemia, lymphoma, fungal infections and hyperproliferative vascular
CC diseases such as restenosis and atherosclerosis. The present sequence is
CC a pUC19-Sep1-5 plasmid protein used in the exemplification of the
CC invention
XX
SQ Sequence 2549 AA;

Query Match   99.9%; Score 9402; DB 5; Length 2549;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 1808; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKDPDPDNPGVINNVLA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKDPDPDNPGVINNVLA 800

Qy   61 TIGELAQVSGLEMRKWVDELFIIMDMLCDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801 TIGELAQVSGLEMRKWVDELFIIMDMLCDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 860

Qy   121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSBS 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSBS 920

Qy   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy   301 FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPEMLRVFMHDNSPGRIVSIKLLAAI 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1041 FWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPEMLRVFMHDNSPGRIVSIKLLAAI 1100

Qy   361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1101 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy   421 RTLDQSPFLRSTAMDTLSSLVFQLGKKYQIFIPMVNKLVLVHRINHQRYDVLICRIVKGY 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1161 RTLDQSPFLRSTAMDTLSSLVFQLGKKYQIFIPMVNKLVLVHRINHQRYDVLICRIVKGY 1220

Qy   481 TLADBEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1221 TLADBEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy   541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy   601 IEIALITSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1341 IEIALITSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy   661 ELEFQKGPTPAILSLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1401 ELEFQKGPTPAILESLIS.INNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy   721 VAYDKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1461 VAYDKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy   781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLGGCQRIVEDWQKILMVR 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLGGCQRIVEDWQKILMVR 1640

Qy   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLGVDPSRQLDHPLPTVHPQVTYAY 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHQELHKLMARCFLKLGEWQLN 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHQELHKLMARCFLKLGEWQLN 1760

Qy 1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVNNFEAVLHYKHQNQARDEKKKLRHA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVNNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy 1081 SGANITNATTAATTAATATTAASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1821 SGANITNATTAA.TTAATATTTAASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880

Qy 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLITWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1881 TVPAVQGFFRSISLSRGNNLQDTLRVLITWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy 1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 2000

Qy 1261 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLBPLHAMMER 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2001 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy 1321 GPQTLKETSPNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS 1380

CC differentiation of a cell or as cell culture additives
XX
SQ Sequence 2549 AA;

Query Match   100.0%;   Score 9413;   DB 2;   Length 2549;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1809; Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALLIKLKDPDPDPNPGVINNVLA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALLIKLKDPDPDPNPGVINNVLA 800

Qy     61 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTIGQLVASTGYVVEPYRK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTIGQLVASTGYVVEPYRK 860

Qy    121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    861 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy    181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy    241 SLGLKCVQFLPQVMPTPLNVIRVCDGAIREFLFQOLGMLVSFVKSHIRPYMDEIVTLMRE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    981 SLGLKCVQFLPQVMPTPLNVIRVCDGAIREFLFQOLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy    301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 QLFGANLDDYLHLHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1100

Qy    361 QLFGANLDDYLHLHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLHLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy    421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy    481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAAARRVSKDDW 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAAARRVSKDDW 1280

Qy    541 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAPVSCWKSELNEDQQDELIRS 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAPVSCWKSELNEDQQDELIRS 1340

Qy    601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy    661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy    721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy    781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQCIDKARDLLDAELTAMAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQCIDKARDLLDAELTAMAG 1580

Qy    841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCGRIVEDWQKILMVR 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCGRIVEDWQKILMVR 1640

Qy    901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy    961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy   1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy   1081 SGANTNATTAANTTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1821 SGANTNATTAATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880

Qy   1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy   1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKNMCE 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKNMCE 2000

Qy   1261 HSNTLVQQAMMVSEELIRVAILWHEMWEGLEEASRLIYPGERNVKGMFEVLEPLHAMMER 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2001 HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLIYPGERNVKGMFEVLEPLHAMMER 2060

Qy   1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVADLTQAWDLYYHVFRRISKQLPQLTS 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVADLTQAWDLYYHVFRRISKQLPQLTS 2120

Qy   1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy   1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy   1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy   1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGRDHPSNLMLDRLSGKILHIDFGD 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGRDHPSNLMLDRLSGKILHIDFGD 2360

Qy   1621 CFEVAMTREKFPEKIPPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2361 CFEVAMTREKFPEKIPPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 2420

Qy   1681 FVIDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy   1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy   1801 CYIGWCPFW 1809
          |||||||||
Db   2541 CYIGWCPFW 2549
```

RESULT 2
AAE13359
ID   AAE13359 standard; protein; 2549 AA.
XX
AC   AAE13359;
XX
DT   12-FEB-2002 (first entry)
XX
DE   pUC19-Sep1-5 plasmid protein.
XX
KW   Rapamycin; FK506 binding protein 12; FKBP12; cytostatic;
KW   transplant rejection; autoimmune disease; lupus; rheumatoid arthritis;
KW   diabetes mellitus; myasthenia gravis; multiple sclerosis; vasotropic;

```
                GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      March  2, 2004, 19:54:47 ; Search time 70 Seconds
                                         (without alignments)
                                         7301.835 Million cell updates/sec

Title:       US-09-517-491-12
Perfect score: 9413
Sequence:    1 LEHSGIGRIKEQSARMLGHL........KQATSHENLCQCYIGWCPFW 1809

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:    1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:      1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database  :  A Geneseq.29Jan04:*
             1:  geneseqp1980s:*
             2:  geneseqp1990s:*
             3:  geneseqp2000s:*
             4:  geneseqp2001s:*
             5:  geneseqp2002s:*
             6:  geneseqp2003as:*
             7:  geneseqp2003bs:*
             8:  geneseqp2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	9413	100.0	2549	2	AAW04235	Human RAP
2	9402	99.9	2549	5	AAE13359	pUCi9-Sep
3	9394	99.8	2549	2	AAR81730	Sirolimus
4	9350	99.3	2549	2	AAW56027	FRAP (tor
5	9341	99.2	2549	7	ADE62170	Rat Prote
6	9325	99.1	2549	6	ABU08621	Rat (rapa
7	9176	97.5	2549	6	ABU61979	Human sir
8	5389.5	57.3	2470	4	ABB61247	Drosophil
9	4690	49.8	1140	5	AAE13360	pGBX-Sep4
10	4685	49.8	1140	6	AAR81731	GST-SEP f
11	4602	48.9	1140	6	ABU61980	Human sir
12	3755.5	39.9	2467	5	ABB91303	Herbicida
13	3440	36.5	1417	7	ADE62172	Human Pro
14	1735	18.4	562	2	AAW04236	Candida R
15	1610	17.1	309	4	AAB84756	FK506 bin
16	1088	11.6	1034	4	ABG21134	Novel hum
17	848	9.0	162	2	AAW04234	Mouse RAP
18	753.5	8.0	2480	2	AAW19724	Cell cycl
19	753.5	8.0	2644	2	AAW13152	Human ata
20	753.5	8.0	2644	2	AAW84271	A human A
21	753.5	8.0	2644	6	ABP97120	Human ATR
22	753.5	8.0	2644	6	AAO30064	Human ATR
23	734.5	7.8	2654	7	ADC27427	Xenopus A
24	726	7.7	3657	6	ABB82900	Human SMG
25	725.5	7.7	2386	2	AAW13153	S. pombe
26	723.5	7.7	2410	2	AAW19723	Cell cycl
27	694	7.4	2162	4	ABB60750	Drosophil
28	693	7.4	2930	4	AAB81485	Human Atr
29	683.5	7.3	2392	4	AAB84758	Human BLI
30	683.5	7.2	870	2	AAW19725	Truncated
31	673.5	7.2	1911	4	AAG67401	Amino aci
32	669.5	7.1	2392	4	AAB84759	Kinase-de
33	663.5	7.0	2354	4	ABB60511	Drosophil
34	639.5	6.8	2157	7	ADD93410	Human lip
35	636.5	6.8	3055	2	AAW19697	ATM mutan
36	636	6.8	2932	2	AAW19675	ATM mutan
37	636	6.8	2989	2	AAW19688	ATM mutan
38	636	6.8	2998	2	AAW19673	ATM mutan
39	636	6.8	2998	2	AAW19691	ATM mutan
40	636	6.8	3026	2	AAW19676	ATM mutan
41	636	6.8	3056	2	AAW19699	ATM mutan
42	636	6.8	3057	2	ABG05614	Novel hum
43	635.5	6.8	3053	2	AAW19694	ATM mutan
44	634	6.7	3054	2	AAW19695	ATM mutan
45	633.5	6.7	3001	2	AAW19670	ATM mutan

ALIGNMENTS

```
RESULT 1
AAW04235
ID   AAW04235 standard; protein; 2549 AA.
XX
AC   AAW04235;
XX
DT   24-NOV-1996  (first entry)
XX
DE   Human RAPT1.
XX
KW   RAPT1; rapamycin binding protein; FKBP; immunosuppressive; fungicide;
KW   anti-mycotic; agonist; antagonist; cell proliferation.
XX
OS   Homo sapiens.
XX
FH                     Location/Qualifiers
FT   Key
FT   Domain             2012..2144
FT                      /label= Rapamycin-binding_domain
XX
PN   WO9533052-A1.
XX
PD   07-DEC-1995.
XX
PF   30-MAY-1995;   95WO-US006722.
XX
PR   27-MAY-1994;   94US-00250795.
PR   20-DEC-1994;   94US-00360144.
XX
PA   (MITO-) MITOTIX INC.
XX
PI   Berlin V,  Chiu MI,  Cottarel G,  Damagnez V;
XX
DR   WPI; 1996-030563/03.
DR   N-PSDB; AAT33872.
XX
PT   Rapamycin binding protein RAPT-1 - used as agonist, or antagonist of
PT   rapamycin cellular proliferation regulation.
XX
PS   Claim 1; Page 72-83; 12pp; English.
XX
CC   Human RAPT1 (AAW04235) is a protein that interacts with the FK506-
CC   binding protein/rapamycin complex and modulates assembly of rapamycin
CC   complexes or cell cycle regulatory proteins. its amino acid sequence was
CC   deduced from a cDNA clone (AAT33872) derived from a B-cell library. The
CC   rapamycin-binding domain of RAPT1 and other RAPT1-like proteins (see also
CC   AAW04235-38) can be used in drug screening assays to identify agents that
CC   modulate the binding of rapamycin binding proteins with FK506-binding
CC   proteins. Such agents can be used to alter the growth and/or
```

```
Qy   1466  NTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGW----------VPHCDTLHALIRDYR 1514
                ||||  :  .  ::  .: :|  |  :| |||  ||:.            |  | :||
Db   1388  NTLLQRNTETRKRKLITICTYKVVPLSQRSGVLEWCTGTVPIGEFLVNNEDGAH---KRYR 1444

Qy   1515  EKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWL-KSPSSEVW 1573
                :  .  ::|   : :: |  |  | ||    : ::  :  |  :
Db   1445  PNDFSAFQCQKKMM-----EVQKKSFEEKYEVFMDVCQNF--QPVFRYFCMEKFLDPAIW 1497

Qy   1574  FDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFPGDCFEVAMTREKFPE 1633
                |:: |  ||||:|  :||||||||||| | :||: |  ||||::  ||  | ||
Db   1498  FEKRLAYTRSVATSSIVGYILGLGDRHVQNILINEQSAELVHIDLGVAFEQGKILPT-PE 1556

Qy   1634  KIPFRLITRMLINAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMD 1693
                :||||||  :  .:  ::|:.|  :  :|||:::|:    :::|  ||||: :|
Db   1557  TVPFRLITRDIVDGMGITGVEGVFRRCCEKTMEVVRNSQETLLTIVEVLLYDPLFDWTM-- 1614

Qy   1694  TNTKGNKRSRTRTDSYSAGQSVBILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEAL 1753
                 :                |: :  :  ::  |      :   :  :.|
Db   1615  -NPLKALYLQQRPE----DETELHPTLNADDQBCKRNLSDIDQS-----------F 1654

Qy   1754  NKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGW 1805
                : | :  ::: |::|| ||             .   : :  :| :.:|
Db   1655  DKVAERVLMRLQEKLKG--VEEGTVLSVGGQVNLLIQQAIDPKNLSRLFPGW 1704
```

Search completed: March 2, 2004, 20:04:05
Job time : 42 secs

```
Qy  1624  VAMTREKFPEKIFPRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVY  1683
               ||  | :|:||||||
Db  2249  EKGKRLPVPEIVPFRLTPNLLDALGIIGTEGTFKKSSEVTLALMRKNEVALMNVIBTIMY  2308

Qy  1684  DPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFI  1743
               |:
Db  2309  D----RNMD--------------------------------HSI-  2316

Qy  1744  GDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYI  1803
                :  ||::   ::|       :
Db  2317  ----QKALKVL---RNKIRGIDPQDGLVLSVAGQTETLIQEATSEDNLSKMYI  2362

Qy  1804  GWCPFW  1809
          ||  ||
Db  2363  GWLPFW  2368
```

RESULT 15

A43100

ataxia telangiectasia-associated protein - human.
N;Contains: 1-phosphatidylinositol 3-kinase (EC 2.7.1.137)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A43100
R;Savitsky, K.; Bar-Shira, A.; Gilad, S.; Ziv, Y.; Vanagaite, L.; Tagle, D.A
.A.; Clines, G.A.; Sartiel, A.; Gatti, R.A.; Chessa, L.; Sanal, O.; Lavin, M.F.; Jasper
, A.; Clines, G.A.; Sartiel, A.; Gatti, R.A.; Chessa, L.; Sanal, O.; Lavin, M.F.; Jasper
Science 268, 1749-1753, 1995
A;Title: A single ataxia telangiectasia gene with a product similar to PI-3 kinase.
A;Reference number: A43100; MUID:95312868; PMID:7792600
A;Accession: A43100
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1708 <SAV>
A;Cross-references: GB:U26455; NID:g9707085, NID:g9707785; PIDN:AAA86520.1; PID:g9870786
C;Genetics:
A;Experimental source: fibroblast
A;Gene: GDB:ATM
A;Cross-references: GDB:593364; OMIM:208900
A;Map position: 11q22.3-11q22.3
A;Note: alternatively spliced transcripts of the single gene may be produced
A;Note: mutations to this gene were found in AT patients from all complementation groups
C;Function:
A;Description: normal protein may function in detection of DNA damage and/or blockage of
C;Superfamily: ataxia telangiectasia-associated protein
C;Keywords: apoptosis; cell cycle control; DNA repair; leukemia; lymphoma; phosphotransf

Query Match   6.5%;   Score 613.5; DB 1; Length 1708;
Best Local Similarity 20.4%; Pred. No. 2.9e-30;
Matches 366; Conservative 282; Mismatches 679; Indels 465; Gaps 72;
```
Qy   246  CVQFLPQVMPFFLNVIRVCGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMREFWV--  303
               |  |        |:          |  :  :|        ::    |:
Db   146  CCDLLSQVCQT--AVTYCKDALENHLHVIVGTLIPLVYEQVE-VQKQVLDDLLKYIVIDN  201

Qy   304  -MNTSIQSTIILL-------------IEQIVALGGEFKLYLPQLIPHMLRVFMHDNSP  348
            : :                     |  || |     |  : :|:
Db   202  KDNENLYITIKLLDPFFDHVVFKDLRITQQKIKYSRGPFSIL--BEINHFLSVSVYDALP  259

Qy   349  -GRIVSIK-ILAAAIQLFGANLDDYLH---------ILLPPIVKLFDAPEAPL--PSRKAA  395
              |:                                  |||   |:
Db   260  LITRLEGLKDLRRQLEHKDQWVTDIMRASQDNPQDGIMVKLVVNLLQLSKMAINHTGBKEV  319

Qy   396  LETVDRLTESLDFTDYASRIIHPIVRTLDQSPELRSTAMDTLSSLVFQLGKKYQ---IPI  452
              |      ||        |                |         |:
Db   320  LRAVGSCLGBVGPIDFSTIAI-------------QHSKDASYTKALKLFE-DKELQWTFIML  367

Qy   453  PMVNKVLVRHRINHQRYDVLICRIVKG--------YTLADEEEDPLIYCHRMLRSCQGDAL  505
                       :                        :          |:
Db   368  TYLNNTIVEDCVKVRSAAVTCLKNILAITKTGHSFWEIYXMTTDPMLAYLQ-----  417

Qy   506  ASGPEVTGPMKKLHVSTINLQXAWGAAARRVSKDD--WLE------WLRRLS------  548
                                              ||   :| |:
```

```
Qy   418  ---PFRTSRKKFLEVPRFDKENPPEGL----DDINLWIPLSENHDIWIKTLTCAFLDSG  469
                                                :
Db   549  ------LELLKDSSSPSLRSCWA------------LAQAYNPMARDLFN---AAFVSC---  585

Qy   470  GTKCEILQLLKPMCEVKTDFCQTVLPYLIHDILLQDTNESWRNLLSTHVQGFFTSCLRHF  529
               |  :|
Db   586  ------WSELNEDQQDELIRSIELALTSQDIAEVTQTLLNLABFMEHSDKGPLPLRDDNG  639

Qy   530  SQTSRSTTPANLDSESEHFPRCCLDKKSQ---------RTMLAVVDYMRRQK---RPSSG  577
               |: ::                           : :  :|
Db   640  IVL--------IGERAAKCRAYAKALHYKE------------LEFQKGPTPAI  672

Qy   578  TIFNDAFWLDLNVLEVAKVAQSCAAHFTALLYABIYADKKSMDDQEKRSIAFEBGSQSTT  637
Db   673  LESLISINNKLQQPEAAAGVLEYAMKHFGELE---------IQATWYEKLIHEWED  718

Qy   638  ISSL-SEKSKEETGISLQDILLELYYRSIGEPDSLYGCGGKMLQPITRLRTYEHEAMWGK  696
          : :  :                                    ||
Db   719  ALVAYDKKMDTNKDDPELMLGRMRCLEALG---------------EWGQLHQQC  757

Qy   697  ALVTYD--LETAIPSSTRQAGIIQALQNLGLCHILSVVLKGLDYENKDWCPELEELHYQ-  753
          |||  |                                          |:
Db   758  CEKWTLVNDETQAKWARMAAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVLALHQDLF  817

Qy   754  ------AAWRNMQWDHCTSVSKEVEGTSYHBSLINNALQSLRDREF  792
                         :
Db   818  SLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEEV----IQYKLVPE  870

Qy   793  STFYESLXYARVKEVEEMCKRSLESVVSLYPTLSRLQAIGELESIGELFSRSVTHRQLSE  852
                                                  ||:
Db   871  ----RREIIRQIWWER-LQGCQR--IVEDWQKILMVRSL  902

Qy   853  VYIKWQKHSQLLKDSDFSFQEPIMALRTVILEILMEKEMDNSQRECIKDILFKHLVELSI  912
             :  |                                    :
Db   903  VVS---PHEDMRTWLKYASL-CGKS------GRLALAHKTLVILLGV--DPSRQLD  946

Qy   913  LARTFKNTQLPERAIFQIKQYNSVSCGVSEWQLEEAQVFWAKKEQSLALSILKQMIKKLD  972
Db   947  HPLPTVHPQVTYAYMKNM-------WKSARKIDAFQHMQCHFVQTMQQQAQHAIATEDQQHKQ  1001

Qy   973  ASCAANNPSLKJTYTECLRVCGNWIAE---TCLENPAVIMQTYLEKAVEVAGNYDGESSD  1029
                                                                   :
Db  1002  ELHKLMARCFLKLGEWQLNLQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAV  1061

Qy  1030  ELRNGKMKAFLSLARP--------SPTQYQRIENYMKSSEFE-------NKQAL  1068
          |:                                              ||
Db  1062  LHYKHQNQARDEKKKLRHASGANINNATTAATTAASTEGSNESEABSTENSPT  1121

Qy  1069  L-----KRAKEEVGLLREHKIQ--TNRYTVKV--------------QRELELDLA-  1103
          :                 :                            |:
Db  1122  PSPLQKKVTEDLSKTLLMYTVPAVOGFFRSISLSRGNNLQDTLRVLTLMFDYGHWPDVNE  1181

Qy  1104  ---LRALKEDRKR-----FLCKAVENYINCL-LSGEEHDMWVFRLCSLWLENSGVSEVNG  1154
                 |:                    ||
Db  1182  ALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRL----IHQLLTDIGRYHPQ--ALI  1233

Qy  1155  MMKRDGMKIPTYKFLPLMYQLAARMGT-KMMGGLGFHEVLNNLISRISMDHPHTLFII  1212
               |:                       :
Db  1234  YPLTVASKSTTTARHNAA--NKILKNMCEHSNTLVQQAMMVSEELIRVALLWHEMWHEGL  1291

Qy  1213  LALANANRDEFLIKPEVARRSRITKNVPKQSSQLDE----------------DRT  1251
           :                                               |||
Db  1292  EEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMERQEWCRKYMKS  1351

Qy  1252  EAANRI--------ICTIRSRRPQMVR--------------SVEALCDAYIIL  1282
                                                    |:
Db  1352  GNVKDLTQAWDLYYHVFRRISKQLP--------QLTSLELQVYSPKLLMCRDELAVPGTYD  1405

Qy  1283  ANL-DATQ-WKT----QRKGINIPADQPITKLKNLE-DVVVPTM---EIKVDHTGEYG  1330
            :                                                     :
Db  1406  PNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVLLKGEEDLRQDERVMQLFGLV  1465

Qy  1331  ---NLVTIQSFKAEFRLAGGVNLPKIIDCVGSDGKERRQLVKGRDDLRQDAVMQQVFQMC  1387
                     ||     ||   |: |
```

```
A;Accession: S47954
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 195-196,'D',188-322;671-744;848-1796;1803-2003;2004-2368 <KA2>
A;Cross-references: EMBL:D11088
R;Weinert, T.A.; Harlow, D.
submitted to the EMBL Data Library, July 1995
A;Reference number: S64650
A;Accession: S64650
A;Molecule type: DNA
A;Residues: 1-715,'P',717-1254,'Q',1256-1275,'G',1277-2368 <WBI>
A;Cross-references: EMBL:U31109; NID:g950172; PIDN:AAA74482.1; PID:g950173
C;Genetics:
A;Gene: SGD:ESR1; MEC1
A;Map position: 2R
C;Cross-references: SGD:S0000340; MIPS:YBR136w
C;Function:
A;Description: required for mitotic cell growth, DNA repair, and meiotic recombination
C;Keywords: ATP; P-loop; purine nucleotide binding; transmembrane protein
F;74-90/Domain: transmembrane #status predicted <TM1>
F;218-234/Domain: transmembrane #status predicted <TM2>
F;771-787/Domain: transmembrane #status predicted <TM3>
F;922-938/Domain: transmembrane #status predicted <TM4>
F;1152-1169/Domain: transmembrane #status predicted <TM5>
F;1288-1304/Domain: transmembrane #status predicted <TM6>
F;1315-1331/Domain: transmembrane #status predicted <TM7>
F;1682-1689/Region: nucleotide-binding motif A (P-loop)
F;1918-1934/Domain: transmembrane #status predicted <TM8>
F;2126-2142/Domain: transmembrane #status predicted <TM9>
F;1688/Binding site: ATP/GTP (Lys) #status predicted

Query Match       6.7%;  Score 628; DB 2;  Length 2368;
Best Local Similarity  20.1%;  Pred. No. 5.9e-31;
Matches  375;  Conservative  305;  Mismatches  614;  Indels  572;  Gaps  77;

Qy    54 VINNVLATIGELAQVSGLEMRKWVD-------ELFIIIMDML-----QDSSLLAKRQVALWT  103
             ||:: : :                                        ::
Db   965 VIDSLIAFI----------LQKWSEFNGKLRNIVYSILDFLIKEKSDLIILKPYTTLAL 1014

Qy   104 LGQ----LVASTGY--VVEYRKYPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDP  156
            :|:      :                 :|: ::
Db  1015 VGKPELGIIARDGQFARMVNKIRSTDLIPIFANNLKSS----------------AS 1053

Qy   157 YKHKVNIGCMIDQSRDASAVSLSESKSSQDSSDYS-----TSEMLVNMGNLPLDEFYPAV  210
            ::|::   : :|:||                       : | :
Db  1054 ----NKYVINQNLDDIEVYL-RRKQTERSIDFTPKKVCQTSDITLVLGAL-LDTSHK-- 1104

Qy   211 SMVALMRIPRDGSLSHHTMVVQAITFIPKSLGLKCVOFLPQVMPTSLNVIRVCDGAIRE  270
              :|::|                   ::|:|
Db  1105 ---FRN----------------LDKDICEKCAK------CISMIGVLDVTKHE 1132

Qy   271 FLFQOLGMLVSFVKSHIRPYMDEIVTLMREFWVMMTSIQSTIILLEQIVVALGGEFKLY  330
            |         ::|:                        ::||
Db  1133 F----KRTTYSENEVYDLNDSVQTIKFLIWVINDI-------- 1163

Qy   331 LPQLIFHMLRVFMHDNSPGR--IVSIKLLAAIQLFGANLD-DYLHLI-LLPPIVKLFDAP  385
            | : :       : :         :||          ::|  ||::|
Db  1164 --LVP----AFWQSNPSKQLFVALVIQESLKYCGLSSESKWDMNHKELYPNEAKLWE-- 1214

Qy   386 EAPLPSRKAALETVDRLITESLDFTDYASRIIHPIVRTLDQSPELRSTAMDTLSSLVFQLG  445
                             :: |:                                |||:  ||
Db  1215 ----------------KFPNSVKTTIYPL--------------LSSL--YLA 1234

Qy   446 KKYQIFIPMVNKVLVLRHRINHQRYDVLICRIVKGYTLADEEEDPLIYQHRMLRSGQGDAL  505
            : :: :|    :  :|: :|  ::                   :                |||
Db  1235 QSWKEYVPL------KYPSNFKBGYKIW--VKKFTL--------------DLL 1266

Qy   506 ASGPVETGPMKKLHVSTINLQKAWGAARR-----VSKD------DWL------  541
            :   :        |:                        :||
Db  1267 KTGTTENHP--LHVFSSLIREDDGSLSNFLLPYISLDIIIKAEKGTPYADIINGIIIEF 1323

Qy   542 EWLRRLSLELLKDSSSPPSLRSCWALAQAYNPMARDLFNAAFVSC--WSELNEDDQQDELIR  599
            |: |            : :                          :  |::|
Db  1324 DSIFTCNLBGMNNLQVDSLRMCY------------------ESIFR-VPEYCKKWATEFKQNYSKLHG 1372

Qy   600 SIELALTSQDIAEVTQTLLNLABFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHY  659
            : : | |                                           |: |
Db  1373 TFIIKDT-----KTTNMLRIDEFLRTTPSD--------LLAQRSLETDSFERSALY 1416

Qy   660 KELEFQKGP-------TPAILELSISINNKLQQPEAAAGVLE-----YAMKHFGELEIQAITW  709
            :|:  ::                 :::::       |||
Db  1417 LEQCYRQNPHDKQNGQLLKNLQITYEEIGDIIDSLDGVLRTFATGNLVSKIEELQYSENW 1476

Qy   710 YEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEBALGEWGQLHQQCCE------K  760
               :|:                               :|||
Db  1477 --KLA-QDCENVLGK---FSDDPKT--TTRMLKSMYD-HQLYSQIISNSSFHSSDGK 1524

Qy   761 WTLVNDETQAKMARMAAAAAWG---LGQWDSMEBYTCMIPRDTHDGAFY--RAVLAL  812
            :|:                       ||:                          :|::|
Db  1525 ISLSPDVKEWYSIGLEAANLEGNVQTLKNW--VEQIESLRNIDDREVLLQYNIAKALIAI 1582

Qy   813 HQDLFSLAQCCIDKARDLL--------DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQY  865
            :::|                                          :| :
Db  1583 SNEDPLRTQKYIHNSFRLIGTNFIITSSKETTLLKKQNLMLHSLYDLSFLSSACDKFEY 1642

Qy   866 K----LVPERREIIRQIWWERLQGCQRIVEDW---QKILMVR---SLVVSPHEDM---R  911
            |          :|:|                    :|::|:|
Db  1643 KSNTTILDYRME---------RIGADFVPNHYILSMRKSFDQLKMNEQADADLGK 1688

Qy   912 TWLKYASLCGKSGRLALAHKTLVLLLGVDPSSIQLDHPLPTVHPQVTYAYMNMWKSARKI  971
            |::|  :|| |                 |||
Db  1689 TFFTLAQLARNNARLDIASESLM----------HCLERRLPQAELEBAEILWKQGEND 1736

Qy   972 DAFQHMQHFVQTMOQQAQHAIATEDDQHKQELHKLMARCFLKLGEWQLNLQGINESTIPK 1031
            || : :   ||                         ::|             ||::
Db  1737 RALKIVQEIHEKYQENS--SVNARDR----------AAVLLKFTEW-LDLS--NNSASEQ 1781

Qy  1032 VLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTA 1091
            ::: |                     :: |        | ||
Db  1782 IIKQYQDIFQIDSKWDKPYYSIG--------LYY--SRLLERKKAEGYITNGRFEVRAIS 1831

Qy  1092 ATTAATATTTASTEGSNSEBAESTENSPTSPLQKKVTEDLSKTLLMYTVPAVQGFFRS 1151
              : |                         |:|::
Db  1832 YFLLAFEKNTA----------------KVRENLPKVITFWLDIAA----AS 1862

Qy  1152 ISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRP 1211
            ||: |  ||:|    : ||                         || :|
Db  1863 ISEAPGNR-KEMLSKAT----------EDICSEVEBALQHCPTYIWYFVLITQLLSGRLLMSHQ 1913

Qy  1212 LVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKQMCEHS---NTIVQQ 1268
            | :|::   | :| :|||                   :|:
Db  1914 SSAQIIMHILLSLAVEYPSHILWYITALVNSNSSKRVLRGKHILEKYRQHSQNPHDLVSS 1973

Qy  1269 AMMVSEELIRVAILWEEMWHEGLEBEASRLIYFGERNVKGMFEVLEPLHAMMERGPQTL-KE 1327
                  :: |: |  :|                                      | :
Db  1974 ALDLTKALTRVC-------------------LQDVKSITSRSGKSLEKD 2003

Qy  1328 TSFNQAYGRDLMEBAQEWCRKYMKSGNVKDITQAWDLJYHVFRRISKQLPQLTSLELQYVS 1387
              |::    :  |                        |::
Db  2004 FKFDMNVAPSAM----------------------VVPVRKNLDI--IS 2027

Qy  1388 PKLLMCRDLBLAVFGTYDENQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVPLLK 1447
            |         |:                         : :||  | :|||:| :|||:|:
Db  2028 P--------LEBSNSMRGYQPFRPVVSLIIRFGSSYKVFSSLKKPKQLNIIGSDGNIYGIMCK 2080

Qy  1448 GHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGVVPHCDTLH 1507
            ::|:||||::::|::  ::  : ::||:|:|:|:||| |:: ::: ||:|: ||| |||:
Db  2081 -KEDVRQDNQYMQFATTMDFLLSKDIASRKRSLGINIYSVLSLREDCGILEMVPNVVTLR 2139

Qy  1508 ALIRDYREKKILINIEHRIMLRMAPDYDHLTLMQKVEVPEHAVNNTAGDDLIAKLL-WL 1565
            ::: :: ::| :|:::   | ||||||:|| | ||:||| ||  : |||::|||   |:
Db  2140 SILSTKYESLKIKVSLK-----SLHDRWQHTAVDGKLEFYMEQV---DKFPPIIYQWF 2189

Qy  1566 --KSPSSEVWFDRTNYTRSLAVMSMVGYILGLGDHPSNLMLDRLSGKILHIDFGDCFE 1623
              || |  :|| ::|||| ||||:||:|||||||:  |:|| || |||||||| |||:
Db  2190 LENFPDPINWFNARNTYARSYAVMAMVGHILGLGDRHCENILLDIQTGKVLHVDF-DCLF 2248
```

```
Qy   1307 MFEVLEPLHAMMERGPQTLKETSPNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYH 1366
               :.|:|
Db   3452 ------------------------EDTNRKVALPREIR----------SVKQL------- 3470

Qy   1367 VFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSK 1426
               : :|||||                 ::       :.  PFFR- GLSDSVTVMNGI
Db   3471 ------------ELVPVVTA----TIPVDRSCQYNE----GSF----PFFR--GLSDSVTVMNGI 3509

Qy   1427 QRPRKLJTLMGSNGHEFVFLLK-GHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRY 1485
               : |||| : |||||::| |||||:::::|.|  |||||
Db   3510 NAPKVVECFGSDGQKYKQLAKSGNDDLRQDAVMEQFFGLVNTFLHNNRDTWKKRLAVRTY 3569

Qy   1486 AVIPLSTNSGLIGWY-----------------------PHCDTLHALIRDY--------- 1513
               :|:|
Db   3570 KVIPFTPSAGVLEWDGTIPLGDYLIGSSRSEGAHGRYGIGNWKYPKC---------REHMS 3622

Qy   1514 --REKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSE 1571
               :|:|                                      HYFFLEK
Db   3623 SAKDKRKAFVDVCTNFRSVM------HYFFLEK-----------------------FLQPAD- 3655

Qy   1572 VWFDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKF 1631
               || ||:|:|:|:::|||::||::::|::|::::|:|:
Db   3656 -WFVKRLAYTRSVAASSMVGYIVGLGDRHAMNILIDQATAEVHIDLGVAFEQGLML-KT 3713

Qy   1632 PEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVWAVLRAFVDPLLNWRL 1691
               | :|| :||:||  :|:|:   |||||::::| :|:::::||:|||::|:| ||| |:|
Db   3714 PERVPFRLJTRDIIDGMGITGVEVGFVFRRCCEETLSVMRINKEALLIVIVEVFIHDPLYKWAL 3773

Qy   1692 MDTNTKGNKSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPE 1751
               : :||::::::  |  | :|||:|:
Db   3774 --SPLKALQRQK-ETEDY---------DGMNL------------------EGLQEEF 3800

Qy   1752 ALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGW 1805
               :|::| : :|||:|::| ||
Db   3801 EGNKDATRALMRVKQKLDG--YEGGEMRSIHGQAQQLIQDAIDTDRLSHMPGW 3852
```

RESULT 14
S46005
ESR1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: MEC1 protein; protein YBR1012; protein YBR136w
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence revision 09-Sep-1994 #text_change 29-Oct-1999
C;Accession: S46005; S45578; S46663; S47954; S64650
R;Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45995
A;Accession: S46005
A;Molecule type: DNA
A;Residues: 1-2368 <BEC>
A;Cross-references: EMBL:Z36005; NID:g536005; PIDN:CAA85094.1; PID:g536430; MIPS:YBR136w
A;Experimental source: strain S288C
R;Becam, A.M.; Cullin, C.; Grzybowska, E.; Lacroute, F.; Nasr, F.; Ozier-Kalogeropoulos,
Yeast 10(Suppl.A), S1-S11, 1994
A;Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 complete
A;Description: An essential gene, ESR1, is required for mitotic cell growth, DNA repair
A;Reference number: S46662
A;Accession: S46578
A;Molecule type: DNA
A;Residues: 1-2368 <BE2>
A;Cross-references: EMBL:X75891; NID:g496856; PIDN:CAA53494.1; PID:g496866
A;Experimental source: strain S288C
R;Kato, R.; Ogawa, H.
submitted to the EMBL Data Library, May 1992
A;Reference number: S46662
A;Accession: S46663
A;Molecule type: DNA
A;Residues: 1-196,'D',198-2368 <KAT>
A;Cross-references: EMBL:D11088; NID:g506874; PIDN:BAA01860.1; PID:d1002337; PID:g506876
R;Kato, R.; Ogawa, H.
Nucleic Acids Res. 22, 3104-3112, 1994
A;Title: An essential gene, ESR1, is required for mitotic cell growth, DNA repair and me
A;Reference number: S47953; MUID:94344772; PMID:8065923

```
Qy    394 AALETVDRLTESLDFTDYASRIIHPIVRTLDQSPELRSTAMDTLSSLVFQLGK------KY 448
               : :|:                   : ::
Db   2514 GILST-ERGQQALSSFDSCBRALIEV---------HGRGVNLDIVEKILLDSQKQFKAEKF 2564

Qy    449 QIFIPMVNKVLVRHRINHQRYDVLICRIVKGYTLADEEDPLIYQHRMLRSGCQDALLASG 508
               ::: :||| ||||| :
Db   2565 SLETPEV------WSTDNKNFDRWICQLV--YCMIALCEDVPI------RLCQNIALLKA 2610

Qy    509 PVETGPMKKLJVS--------TINL-----QKAWGAARRVSKDDWLEWLRRLSLELL 552
               :::                |||||::|
Db   2611 EISELLFPSVVVSLAGRIGMDINLHDLITSQVKEHIFTDSNKLTKS------KQVMLNTL 2664

Qy    553 KDSSSPSLRSCWALACAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRSIELALTSQDIAE 612
               :.: : :||
Db   2665 NE----LRMCYVL-------ERSIFS---------GQTKREKNSRSCSTAAKIRDV-- 2700

Qy    613 VTQCTLLNLAEFMEHSDKGPLPLRDDNG-------------IVLLGERAAKCRAYAKALHYKEL 662
               :: ::
Db   2701 -----------ESGSNGMAASITTNWEKVYWLSIDYLVVAGS-AVVCGAYLTASMVYEY 2747

Qy    663 EFQK------------GPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELRIQA 707
               ::
Db   2748 WCEEKFGNLSLGDPDFSYHDKLPDHVEILVSAITRINEPDSLYGVI-----HSNKLSAQI 2802

Qy    708 TWYEKLHEWEDALVAYD------------KRMDTNKDDP-------------EL 736
               :.:                  : ::
Db   2803 ITFFEHEGNWTRALEYDLQARSQKMVVPSSLSENLEVEQFQPTTSARHSVFGEGEVQRQP 2862

Qy    737 MLGRMRCLEALGEWGQLHQQC-------CEKWTLVNDETQAKMARMAAAAWGLGGWDS 788
               : :                     ::|
Db   2863 FKGLIRSLQQQTGCMHVLDLYCRGLTSREGCFQY------DPEFIELQYEAAWRAGKWD- 2914

Qy    789 MEEYTCMIPRDTHDGAFYRAV-----LALHQDLFSLAQQCID----KARDL--LDAELT 836
               : :|                                         : :|
Db   2915 -----FSLLYPQ-THCQPLQHAKNNYHESLHCCLRALQEGDYGFYGKLKDTKKELVLSIS 2970

Qy    837 AMAGESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIR------------QI 878
               :                                         ::
Db   2971 RASEESTEFIYSTVVKLQILHHIGLVWDLRWTSSHQSVHGYLVKQMACVDPVIPTMDQL 3030

Qy    879 WWERLQGCQRIVEDWQKILMVRSL---VVSPHEDMRTWLKYASHLCGKSGRLALAHKTLVL 935
               :: |                 :: ::|
Db   3031 SW-------LNKDWNSIITCQTOLHWTLLEPFIAFRRVLLQILQGCEKTMQHLLQSASLL 3082

Qy    936 LLGVDPSRQLDHPLLPTVHPQVTYAYMKNMWKSARKIDAPQHMQHFVQTMQQQAQHAIATE 995
               : :
Db   3083 RKGT------RFSHAAASLH--EFKFLCARSNGQQPVPDWLGKLEEAKLLHAQGRHEVSIS 3135

Qy    996 DQQ---HKQELHKLMARCPLKLGEWQLNLQGINESTIPKVLQYY------SAATE------ 1041
               :
Db   3136 LANYILHNYQLKEEASDIYRVIGKWLAEYRSSNSRTI---LEKYLRPAVSLAEEQSSKIC 3192

Qy   1042 ------HDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTA 1095
               :
Db   3193 KRLVDRQSQTWF------HLAHYADALFKSYBERLSSSE--WQAALRLRKH 3235

Qy   1096 ATATTTASTEGSNSESERAESTENSFTPSPLQKKVTEDLSKTLLM------YTVPAVQGFF 1149
               :
Db   3236 KTKELEVFIKRFKSSKKAEQSDYSLKIQDLQKQLTMDKEEABKLQVDRDNFLKLALEGVK 3295

Qy   1150 RSISLSRGNNLQDTLRVLIWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTP 1209
               :
Db   3296 RCLEIGDKYDVRVVFRQVSMWFSLASQKNVIDNMLSTIKEVQSYKFIPLVYQIASRLGSS 3355

Qy   1210 RPLVG-------RLIHQLLTDIGRYHPQALLYPLTVAS-----KSTTTARHN------- 1249
               :
Db   3356 KDBSGSNSFQSALVSLIRKMAID----HPYHTLQLLALANGDRIKDNQRSRSNSFVDMD 3411

Qy   1250 ---AANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKG 1306
               : |::|.
Db   3412 KKLAAEHLLQDVSHYHGMIRQMKQLVDIYIKLAEL------------ETRR------- 3451
```

Qy   850 MVSCHMLSELEVIQYKLVPERREIIRQIW----WERLQGCCRIVEDWQKILMVRSLVVS 905
         ::|:|:: ::::::| : : :|: |:::| :::|
Db   446 LIIANLVNKMTGVSQWKNKLTDYEIFDRNEEGNDGDKLAICRKLMHWGRHTKSNRGQSCA 505

Qy   906 PHEDMRTWLKYASLCGKSGRLALAH--------KTLVLLIGVDPSRQLDHPLPTVHPQVTY 958
         |||| :: ::|:: |
Db   506 AHSEIIRLSRKTSNC---ELAFFHINSALRGEKLAAWQRLEVERQRLKLVKTQNLDVRI 561

Qy   959 AYMQNMWKSARKI---------DAFQEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARC 1010
         ::| :::|| |::   :::|| |: | : :
Db   562 REMNEVFGSLAEVFTTSCQLKSDFQ---MVDGMIKEKMISEGYNEDIAKREEH--MSRA 615

Qy 1011 FLKLGEWQLNLQGINESTIPKVL---------------------QYYSAATEHDRSW 1046
         ::|::|| : |:|
Db   616 SIQLADFFQSLPELENVLAPNLPPTIIWSELQRSDSLSAGYHGIVGSLFHIASEMCPSL 675

Qy 1047 YKA-----WHAWA-----VMNF-EAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAA 1096
                             : | :: :::|:::|:: :
Db   676 AKAHLKARWAYEIAKIKNFPAENLSFYKPGKDETENEELLKSLRATSLVNLEKLVRAAI 735

Qy 1097 TATTTASTEGSNSEBAASTENSPTPSP---LQKKVTEDLSKTLLMYTVPAVQGFFRSI 1152
         :   :::|| |: :   : |:|||
Db   736 SDDLRAN-----------NILAPNSHFMHIWKMVRDHRTKFLSI----AVTSYFQFI 777

Qy 1153 SLSRG--NNLQ-------DTLRVLTLWFDYGHEWPDV-NEALVEGVKAIQIDTWLQVIP 1200
         : :|: ||:|: | :|::| :: :: |||::
Db   778 QNMSGDCDNLPYSKKEETTLATLRIILELLVKHG---DVLIDVINDGLNKTNVHHINKEILP 834

Qy 1201 QLIARIDTPRPLVCGRLIHQLLTDIGRYHPQALLYPLTVASKSTT-------ARHN------ 1249
         ||: ||||: |:   : : ||| :| :|: : |
Db   835 QLFARLSHPSEHIRKTLWTVDLISKICTAAPHAVVFQVVSGAASSTDGBELEEQQNDDRNR 894

Qy 1250 ---AANKILKNMCEHSNTLVQQAMMVSBELIRVAILWHEMV-----------HE------- 1289
             :|::| :: : |:| ||:|: |:
Db   895 VRACCEKLETNMSQSYPNLVKDVRQFVAELERINLLNEBKWSVVMGTMEHEMEKRLSLIR 954

Qy 1290 ----GLEEASRLYFGERN-----------VKGMFEVLEPLHAM-MERGPQTLKETSFNQAY 1334
              ::| : | :::|: ||:| :   :
Db   955 TENAKTESALHLTASVKNDIIVKRTQLLTRQIIFDVLDELYQQTVIEPPKSKNEEFVTAF 1014

Qy 1335 GRDLMEBAQEBWCRKYMKSGNVKDLTQAWDLYYHVFRRISKGLPQLTSLELQVVSPKLLMCR 1394
         ||:|:|: |: :|:||| ::: | :|||:::| ::|:: :: ::|||:|:|||
Db 1015 AEVLTNAFQESRISRTTSPEKSWIPFKNLIANFVHRNSKK--GMQTFETEDISPYLIASLS 1072

Qy 1395 DLELAVPGTYDPN-QPIIRIQSIAPSLQVITSKQRPRKLTIMGSNGHEPFVFLLKGHEDLR 1453
         |:|:| :|::: |::| |::|| |::| |||:|::|:||:||||:|:||:||||||:
Db 1073 NSCVPMPGQEBSVEFDRVVSISRVARQVTILPFKTRPKKLGFVGSDGKQVAFLFKGREDLH 1132

Qy 1454 QDERVMQLFGLVNTLLANDPTSLRKNLS---IQRYAVIPLSTNSGLIGWPHCDTLHAJI 1510
         ||||:||:||:|||||||||:|:|:|:   ::|:||:|:|:||:|||||||||:|:|:
Db 1133 LDERVMQFLRLCNVMLQPGKGHRQSVAAYQAHYAVIPLGPRSGLLIKWVEGATPMFHIY 1192

Qy 1511 RDYRKKKIL--------------LNIEHRIMLRMAPDYDHLITLMQKV----------- 1544
         |:|:||::   ::|::::|:|::|||:|| ||:|||
Db 1193 RKWQOMKEKALKQATKNGETVPEIERPSNMYHN-MIRLA-FADH------KIDSSITSDRS 1245

Qy 1545 ---------EVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMGYILGL 1596
                  :|:| :| |||:|::|:|: |:|||:|| :|||||:|:||||||||:|||
Db 1246 KWPAEILEEVFESLITAKTPTDLISRELWMRANDATTWWSVTKRYSRSLAVMSMWGSVLGL 1305

Qy 1597 GDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPEKIPFRLTRMLTNAMEVTGLDGNY 1656
         ||||:|:|:|||:|::|:||||| |::| ||:||:||| |:||:| :|||:|:|||
Db 1306 GDRHLDNLLVDLKWGHVVHIDYNICFDKGKNL-RIPETVPFRLTRNMRHALGPSEMYGTF 1364

Qy 1657 RITCHTVMEVLREHKDSVMAVLBAFVYDPLLNWRLMD-TNTKG------------NK 1700
         |::| : |: :|||:| :| |:||||:||:| |::|:   ::
Db 1365 RESCVHVLSTLRSGHQVLIMLLDAFVFDPLVDWTSHEHTATSGVSLALQLAVYGSNKKTX 1424

Qy 1701 RSRTRTDS-------------YSAGQSVE--ILDGVELGRPA-------HKK 1730
         |||| |:   : | ||: ::|||:||:|| |||
Db 1425 AKERILTDAMELLNLRMSEVQTLMLANRDDLLHWMKQVTECLLIENSMLGANAIYAQQRVK 1484

RESULT 13
T51174

ataxia-telangiectasia mutated protein (Atm) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 02-Sep-2000
C;Accession: T51174; T13008
R;Garcia, V.; Salanoubat, M.; Choisne, N.; Tissier, A.F.
Nucleic Acids Res. 28, 1692-1699, 2000
A;Title: An ATM homologue from Arabidopsis thaliana: complete genomic organisation and
A;Reference number: Z25324
A;Accession: T51174
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3856 <GAR>
A;Cross-references: EMBL:AJ250248; PIDN:CAB86487.1
A;Experimental source: cultivar Columbia
R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; S
submitted to the Protein Sequence Database, July 1999
A;Reference number: Z17586
A;Accession: T13008
A;Status: significant sequence differences
A;Molecule type: mRNA
C;Genetics:
A;Gene: atm

Query Match   6.8%;   Score 641.5; DB 2; Length 3856;
Best Local Similarity   19.5%;   Pred. No. 1.9e-31;
Matches   409; Conservative 326; Mismatches 748; Indels 611; Gaps   87;

Qy   9 IKEBQSARMLGHIVSN-APRLIRPVMEPILKALILKLKDEPDENPGVINNVLATIGELAQ 67
         |::| |::|:|::|| ||:|:||:|: |:||:|:|||| |:|:| |::|:|||| |:|
Db 2073 VLQNSILYVGETSENERDKLIKQNMVSIV-SPILSCASSPEPPVPTFSRDTISLAVQTV 2131

Qy   68 VSG-LEMRKW-------------------VDELFIIMDMLQDSS----------LLAKRQV 99
         ||| ::|:| ||:|: ||::|:: ||:|:
Db 2132 VDGFLENTDYPKNAAITDRINIFRPDRVFMFITEMHYRMSAACHIRHTRHHLAALEELTI 2191

Qy   100 ALWTLGQLVASTGYV--VEPYRKYPTLLE--------VLLNFLKTEQNQGTRREAIRVLG 149
         |:|||||:|:|:|: ::| |:||:|: |::|:|||:|| ||| |:|:||
Db 2192 LLIGHRALVPSSLNYIFNLVQQFIGYPSLQDQCCSIASCLLDLFKS---NPAKEIVSVLG 2247

Qy   150 LLGALDPYKHKVNIGMDLQSRDASAVELSESKSSQ------------DSSDYSTSEMLVNMGN 200
         |||| | |:|::| ||||:| ::|: |::||| | |:| ::||:|:|:
Db 2248 --DQIQFLVSK.VTCCIDABAD---TKISGAKSSQLVNLLHKLVVSSDSLNEDIRDLEP 2302

Qy   201 LPLDEFYPAVSWVALMRIFRDQSLSHHTMVVQAITFI----------FKSIGLKCVQ 248
         ||||:| |:|: |:|:|||| ::||:|||:||:|| |||::|:|:|
Db 2303 LP-DLKFFQVIRESHIRICEAYS-PRNHLKVEHSTFLIYIFLEILSLSNFLFLSCCTIQ 2360

Qy   249 ---------FLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMREF 301
                  ||||:|| | ||::||||:||:| |:||:|:|||||::|:||||:||||||
Db 2361 QCSRRSNVLP---PRFLS------RSLQALHNKLIASEVSQE------DTNGETAETF 2403

Qy   302 WVMNTSIQSTILLIEQIVVALGGGEFKLYLPQLI--------PHMLRVFMHDNSPGRIVSI 354
         |||||:::|| |:|::|:| |||||||:|||: |:|:|:| |||||:|:|:
Db 2404 WQSDDEIVNAVWTLVRVSASDEADSMRLLVSDFLSRIGIRDPHTV-VF---HLPGNLVS- 2458

Qy   355 KLLAAIQLFGAN------------------LDDYLHLLLPPIVKLFDAPPAPLPSRK 393
         |::||||||| ::|| |:|:|:|||:|| |: |||
Db 2459 --MEIGLQGFGHNTGSKVRSLITENGISDETLITLLNFLKKYLLDDSVKIIDVISQTL--R 2513

Qy 1731 TGTTVPBSI--HSP---------------------DGLVK-------IG------- 1749
         |::| :   : :| ||:|::   :|
Db 1485 AGTELRAVTRHHALAKELRPLIRVIGKEREBFADYL.KFYKQALFDPLLKGHSALRNELD 1544

Qy 1750 ---PEAL------- 1753
         |||
Db 1545 IDTCVYNFNIVMQNIDNVFGALVNLSFTPIETITSRTSQQFKPPGLENVWVKDQQE 1604

Qy 1754 NKKAIQIINRVRDKLTG--RDFSHDDTLDVPTQVELLIKQATSHENL.CQCVIGW 1805
         |::|::::|::|:|:: |:   :: |:|||:|:|::|:|:|:|: |:|:::|
Db 1605 NSQAREVVRRVERRINGWLDSGAGPDRKLSPREBADILIAEATSTPNLSQMYBGW 1659

```
Qy    387  APLP--SRKAALETVDRLTESLDF--------TDYASRIIHPIVRTLDQSPELRSTAMDTLS  438
               ||     ||||          :||      |:      ::            ::  |||
Db   1067  LPLNILSDDFAVLALTSLCRGYQFQQNTKHVDSFSLAIQETLAICGISPK-EQKKVQLWQ  1125

Qy    439  SLVFQLGKKYQIFIPMVN----------KVLVRHRINHQRYDVLICRIVKGYTLAD  484
               :|     |:|:                 |  :       :
Db   1126  SLP--ARMRQLMEPMLHSCYTCVHRPSTCLQQPLFGSHYSNYY-----------  1167

Qy    485  EEEDPLIYQHRML-----RSGQGDALA--------SGPVET-GPMKKLEVSTINLQKAWG  530
               |:|||:                  ||          ||    |     |  | ||||
Db   1168  -EEWAFLWASRLIDYLPSSGKRHLLSSYKPCIKRDSNMLSTPYPYILLHA----LLECTT  1222

Qy    531  AARRVSKDDWLEWLRRISLELLKDSSPSLRSCWALA-----QAYNPMARDLFNAAFVSCW  586
               ||:  |                 ::         :: :       :
Db   1223  EQRNHIQEEFMAVLQ------ANEESSSVRGRQELGAIKENAFKQFESRKYAAGIKPLA  1276

Qy    587  SELNEDQQDELIRSIEHLA--LTSQDIAEVTQTLNLAEFMEHSDKGPLPLRDDNGI-----  640
               |:|:||                :       :|     |::    | :|
Db   1277  STLVSDRKEDSSRVPRIAGKLCAELLDFLQRWLREWQRIHGRSTGGKPETIDSNYRKIH  1336

Qy    641  ----------VLLGERAAKCRAYAKALHYKELEFQKG--PTPAILES---LISINNKLQQPE  687
                          |||              ::||        |:      ::||| |||
Db   1337  EFLNLIPKLLVSRASYNCGEYARALSYLESYLEEGEDKSQRLLEQFTFLVEVYGSLRDPD  1396

Qy    688  AAAGVLE--YAMKHFGELEIQATWYEKLHEWEDALVAYDKKMDTN----------  730
               |||||     ||||||                      ::
Db   1397  SVEGAVQVRSYDMSVTQDILVN----RLVERQQDMITSYEQLLSSTDQMQPDHVRAMIDA  1452

Qy    731  --KDDPELMLGRMRCLEALGEWGQL----HQQCCEKWTLVNDETQAKVARMAAAAAWGLG  784
                |||       |  ||| |||        |: | |                  ||
Db   1453  YLRDTP-----KTAQLIADGLWQRLSDRYSDQC----------FABCKSELLWRLG  1493

Qy    785  QWDSMEERYTCMIPRDTHDGAF-YRAVLALHQDLFSLAQQCIDKARDLLDAEL---TAMAG  840
               |  ||||:||            |:|| |  ||             ||  |::      |
Db   1494  SYDEMEELQSNWPAQCSQGCLKLRRPLTTRIEFDSL----LDGMRESVLEELRSCSAVQQ  1549

Qy    841  ESYSRAYGAMVSCHMLSEJ---EEVIQYKLVPERREIRQ----------IWWERLQSCQRI  889
               ||: |  :|          :  ||:  | | ||  |:              ||||| ||:
Db   1550  HSYANAYDAVLKLHLVHELHCSQELVE-KLEQDRDEDNQEKLMKNYFDDWQYRLQIVQFQ  1608

Qy    890  VEDWQKILMVRSLVVS------------PH---EDMRTWLKYASLCKSGRLALAHKTL  933
               :: |     | :  |              |    :  | :| |    |    |   |
Db   1609  VRIQESIYSPRRNIIAELQRRLTDRNHLLPHLKTELARIWLNSAQINRNAGQLQRAQLYI  1668

Qy    934  VLLLGVDPSRQLDHLPLPTVHPQVTTAYMKNMWKSARKIDAFQHMQHFVCTMQQQAQHAI-  992
               :||  ||        |::|||| |:          |   | :: |     :  |
Db   1669  LXAAEYQPS----------GLFIERAKLLWQKGDQVMAMNYLEEQLSIMRSGCCGNVK  1716

Qy    993  -ATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQ-YYSAATEHDRSMYKAW  1050
               || |||  |  |||    |||   |  :     |        ::|:  |
Db   1717  QLAAEQRH---------------LFFRG-------KYLQAVYSAESMH-----------  1742

Qy   1051  HAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATPATTTASTEGSNSE  1110
               |   | :|      |     :    |  ||:|:   ||||  ||: |   || |||
Db   1743  ----LCADAVLKYFQEAIAVHRQSESCHVQMAQFLEKILEA---------RQGGKSE  1786

Qy   1111  SEAESTENSPTPSPLQKKVTEDLSKTLLMYTPAVQGFFRSISLSRGNNLQDTLRVLTLW  1170
               | ||:       ||   |    ||             ||| | ||::|  |: ::||:||
Db   1787  PTGEQDD----------------MLINVMVNYAKSLRVGSEHVYQSMPRLLISIW  1824

Qy   1171  FD-------YGHWPDVNEALWBGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQLLTD  1223
                |         |||         |:  ||:||:|:| |  |  :|: |:|||: :||  |
Db   1825  LDTTESSTNTEQVXKXMNDLLTNCCTALPTAVFYTTVSQMLSRLCHPVNDVFTVLRNVIIK  1884

Qy   1224  IGRYHPQALLYPLTVASKSTTTARHNAANKILKNMCEHSNT---LVQQAMVSEBLIRVA  1280
               || :   |  | |||||:          | :|||::            ::  | ||
Db   1885  LVEAYPQQSLWMLLPHFKSAKAHRIKRCKLVLTDSRLQNSTFQKLLQDFNSLTERLMDI-  1943

Qy   1281  ILMHEWWHEGLEEASRLYFGFRNVKGMFPEVLEPLHAMMERGPQTLKETSFNQAYGRDLME  1340
               |:| ::| |||  |:  :|:   ||       :|: |     |       ||
Db   1944  -----FRSVELLTKFEKLNNTVNEKRQVVD-----------WSARERPQF  395

Qy   1341  AQEWCRKYMKSGNVKDLTQAWDLIYHVFRRISK---QLPQLTSLEL----QYVSPKLLMCRD  1395
               | |      ||||| |   ||:||||| |:|:      ::                ::
Db    396  V-----ESAFSQTM--RRRVSRVTELLDIQKDFSAMGALALSADSCKLYSDISSTS  445
```

```
Db   1955  --------KLSLDLD----------TRLSKLCKQPEFSQILLPFEKYMQPTLPLNSD  1992
```

```
Qy   1396  LELAVPGTYDNQPIIR------IQSIAPSLQVITSKQRPRKLTLMGSNGHEFVPLLK  1447
               | |   ||| |:||         :||   :|||  : |  :| | | |||  |  ||
Db   1993  SNSSEGSHLPANQSTVNWFPYQQIYISGFQESVLILRSAAKPKKLTIRCSDGKDYDVLVK  2052
```

```
Qy   1448  GHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLH  1507
               ||:||:|  ::| ||||| ||| |::| | || :|| |||  | ::|| || |  | |
Db   2053  PKDDLRRDARLMEFNGLVKRYLHQDARPQRRLHIRTYAVLPFNEBCGLVEWLPNLASYR  2112
```

```
Qy   1508  ALIRDYREKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLL------  1563
               :  :  | || :   :     |          || |: |  |    | :
Db   2113  SICMNLYAQRRIVMSTRQLQSLAVPL--HESIRRKREVF--------------  2157
```

```
Qy   1564  -WLK----SPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKI  1613
                |||      | | | ||  :  ||| ||:| | |||| |  || |:  | ||:
Db   2158  PPVFQEWLRQRFATPHS--WYEARNTYIRTVAVMSMVGYILGLGDRHGENILPAEGNGDA  2215
```

```
Qy   1614  LHIDFGDCFEVAMTREKPPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDS  1673
               :|||||| |: |||:| || | |: | :| | |:|| :|| ||  || |  | :||:|
Db   2216  VHVDFNCLFNQGELL-PYPEVVPFRLTHMIVAMGPLGVEGSFRKCCEITLRLLKQESKT  2274
```

```
Qy   1674  VMAVLBAFVYDPLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGT  1733
               : || |||:|                      |                | |  |
Db   2275  LMSILRPFVYD-------------------------------------VGAQTRKGAAT  2296
```

```
Qy   1734  TVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTG--RDFSHDDTLDVPTQVELLIK  1790
               |: |: |  |||| |  ||:|||::|:|  || :||       | |||| || ||:|||
Db   2297  A-------------------------KITKDVQRIADRLQGHVKRQQANSIPLSTEGQVNFLIN  2335
```

```
Qy   1791  QATSHENLCQCYIGWCPF  1808
               || |||| || ||| ||:
Db   2336  EATKVDNLASMYIGWGAF  2353
```

RESULT 12
T28923
hypothetical protein C48B6.6. - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28923
R;Minx, P.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid C48B6.
A;Reference number: Z20544
A;Accession: T28923
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1663 <MIN>
A;Cross-references: EMBL:U97189; PIDN:AAC48166.1; GSPDB:GN000019; CESP:C48B6.6
A;Experimental source: strain Bristol N2; clone C48B6
C;Genetics:
A;Gene: CESP:C48B6.6
A;Map position: 1
A;Introns: 43/2; 111/1; 185/3; 221/3; 298/2; 327/3; 364/2; 411/2; 500/1; 595/1; 643/3;
2; 1594/1; 1638/3

Query Match   6.9%; Score 645; DB 2; Length 1663;
Best Local Similarity   21.2%; Pred. No. 2.7e-32;
Matches 304; Conservative 223; Mismatches 502; Indels 406; Gaps 50;

```
Qy    675  SLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDP  734
               :::  | ||                :||| |
Db    327  AMVDIINRIRLPQATD--YYKVVLFGE------------ESNDSTITED----------  361
```

```
Qy    735  EIMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKVARMAAAAAWGLGQWDSMEBYTC  794
                   :                                     |
Db    362  ----FRSVELLTKFEKLNNTVNEKRQVVD-----------WSARERPQF  395
```

```
Qy    795  MIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAGESYSRAYG------A  849
               :                   |
Db    396  V-----ESAFSQTM--RRRVSRVTELLDIQKDFSAMGALALSADSCKLYSDISSTS  445
```

Db   2166 LDFA------LRSPNPGDIFEKKI-------LPKPFPVFYFWFVESFPBPNNWVTSRQNYC 2213

Qy   1582 RSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLTR 1641

Db   2214 RTLAVMSIVGYVLGLGDRHGENILFDFTGEALHVDFNCLFDKGLTFEK-PEKVPFRLTH 2272

Qy   1642 MLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKR 1701

Db   2273 NMVDAMGPTGYEGGFRKASETMRLLRSNQDTLMSVLESFLHDPLVEWN--------- 2321

Qy   1702 SRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQII 1761

Db   2322 --------------------------RKKSSSKYP--------NNEANEVL 2338

Qy   1762 NRVRDKLTGRDFSHDDT--LDVPTQVELLIKQATSEHNLCQCYIGWCPFW 1809

Db   2339 DIIRKKPQG--FMPGETIPLSIEGQIQELIKSAVNPRNLVEMYIGWAAYF 2386

RESULT 11

T13288

mei-41 protein - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 13-Aug-1999 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C;Accession: T13288

R;Hari, K.L.; Santerre, A.; Sekelsky, J.J.; McKim, K.S.; Boyd, J.B.; Hawley, R.S.

Cell 82, 815-821, 1995

A;Title: The mei-41 gene of D. melanogaster is a structural and functional homolog of t

A;Reference number: Z11072; MUID:95401271; PMID:7671309

A;Accession: T13288

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-2354 <HAR>

A;Cross-references: EMBL:U34925; NID:g998351; PIDN:AAC46881.1

C;Genetics:

A;Gene: mei-41

A;Cross-references: FlyBase:FBgn0004367

A;Introns: 650/3; 748/3; 2313/3

C;Function:

A;Description: involved in cell cycle checkpoint and meiotic recombination

Query Match   7.0%;   Score 663.5;   DB 2;   Length 2354;
Best Local Similarity   19.4%;   Pred. No. 3.2e-33;
Matches 387; Conservative 316; Mismatches 706; Indels 589; Gaps 74;

Qy   9 IKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLATIGEL-AQ 67

Db   747 VKQTVAELLIYFNRNPTFVMRSFQS------LLQLS------IGSLEELLSSQ 786

Qy   68 VSGLEMRKWVDELFIIIMDMLQDSSLLAK--------RQVALWTLGQLVASTGTVVVEPYRK 120

Db   787 TANAEFANFIAERFLGVITYFE--SCLSEPSFEKPLKEETLYSLGQIMRFVG--SQHVTQ 842

Qy   121 YPTTLEVLLNFLKTEQNQGTRREAIRVLGLLGALLDPYHGKVNIGMIDQQRDASAVS---- 176

Db   843 FRFKIIAMLSFVHTLQEPRLQRICLKIWHI-------FLHVVNVQELGPSLGRIVATILQPL 896

Qy   177 LSESKSSQDSSDYSTSEMLVN--MGNLPLDEFY-------PAVSMVALMRIFRDQSLS 225

Db   897 LADNESVKQVNDLYBFIILRNASMLGTFITDLYFLDRMENVSPSI-------QKCI 945

Qy   226 HHHTM--------VVQATTFIFKSLGLKCVQFLPQVMPTFLNVIRVCDGAIR 269

Db   946 RRHTAHLDLGLAEEEDQSPPLVDQLRF-QKHTDECLQ---------VR 986

Qy   270 EFLFQQLGMLVSFVKSHIRPYMDEIVTLMREFWVMNTSIQSTIII--LIEQIV-VALGGE 326

Db   987 VYALQHLGDLF---GRRRP-------------KLNSTILSELPLEPMLEQIVNVLMAG-- 1028

Qy   327 FKLYLPQLIPHVLRVFVMHDNSPGRIVSIKLLAAIQLPGANLDDYLHLLPPIVKLPDAPE 386

Db   1029 --------------------CQHDDSQLQMASAKCLGELGAIDAS------YLPSNYNFASPQH 1066

Db   1290 ELEVNKVIEEFQLVINQPGPDGLNSVG--QQRYTSFVDVFFKIVDYLNKWLRMRKKRNWD 1347

Qy   593 QQDELIRSIELALTSQDIAEVTQTLLNIAEFMEHSDKGPLPLRDDNGIVLLGRAAKCRA 652

Db   1348 RRSAIARKENRYMSVEDATSRESSISKVESFLSRFPBSKTL------GIVSL------NCGF 1396

Qy   653 YAKALHYKELEFQKGPTP-AILES----LISINNKLQQPEAAAGV-LEYAMCHFGELEIQ 706

Db   1397 HARLFWEQHIRNATAPYAALESDYIRVLQBEIVAGIDDPDEIEAVSLNF---HDYSFDQQ 1453

Qy   707 ATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCBKWTLVND 766

Db   1454 LLLHENSGTWDSALSCYEIIIQKDPENKKAKIGLLNSMLQSGHYESL-VLSLDSFIINDN 1512

Qy   767 ETQAKARMAAAAWGLGWDSMEEYTCMIPRDTHD------GAFYRAVLALHQDLFSLA 820

Db   1513 HEYSKMLNLGIEAGWRSLSIDSLKK--CLSKSNLESFEAKLGSIFYQ--YLRKDSFA- 1565

Qy   821 QQCIDKARDLLDAELTAMAGESYSRAYGAMVSCH-MLSELEEVIQYKLVPERREIIRQIW 879

Db   1566 -ELTERIQPLYVDAATAIANTGAHSAY---DCYDIISKLHAINDFSRIAE-------- 1611

Qy   880 WERIQGCQRIVEDWQKIMVRSL---------VVSPH----------EDMRTWLKY 916

Db   1612 ------TDGIVSDNLDIVLRRRISQVAPYGKFKHQIILSTHLVGYEKFENTKKTAEIYLEI 1665

Qy   917 ASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNWWKSARKIDAFQH 976

Db   1666 ARISRKNGGFQFQRAFNAII------KAMDLLDKPLATIE------HAQWWHQGQHRKAISE 1713

Qy   977 MQHFVQTMQQQAQHAIATEDQQHKQEL-HKLMARCFLKLGEW------QLNLQGINESTIP 1030

Db   1714 LNFSLNNNMFDLVDEHERPKNRKETLGNPLKGKVFLKLFKWLGKAGQLGLKDL-ETYYH 1772

Qy   1031 KVLQIYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATT 1090

Db   1773 KAVEIYS--ECENTHYYLGH------HRVLMYEEEQKLPVNBQSERFLSGELVTR--- 1819

Qy   1091 AATTATATTAASTEGSNSESEABSTENSPTPSPLQKKVTEDLSKTLLMYTVVPAVQGFFR 1150

Db   1820 ----------------------IINEFGR 1826

Qy   1151 SISLSRGRNLQDTLRVLTLWPDYG---------HWPDVNBALVEGV-----KAIQIDT--- 1194

Db   1827 SLYYGTNHIYESMFKLLTLWLDFGABELRLSKDDGBKYFREHIISSRKKSLELMNSNVCR 1886

Qy   1195 ---------WLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTT 1245

Db   1887 LSMKIPQYFFLVALSQMISRVCHPNNKVYKILEHIIANVVASYPGETLWQLMATIKSTSQ 1946

Qy   1246 ARHNAANKILKNMCEHS---------NTLVQQAMMVSBELIRVAILWEEWWHEGLEEA 1294

Db   1947 KRSLRGKSIlNVL--HSRKLSMSSKVDIKALSQSAILITEKLINL---------- 1989

Qy   1295 SRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEBAQEWCRKYMKSGNV 1354

Db   1990 -----------------------CNTRINSKSV 1999

Qy   1355 KDLTQAWDLYYHVFR-----RISKQLFQLTSLELQYVSPKLLMCRDLBELAVPGTYDPNQP 1409

Db   2000 K------MSLKDH-FRLSFDDPVDLVIPAKSFLDI------TLPAKDANRASHYPFPKTQP 2047

Qy   1410 IIRIQSIAPSLQVITSKQRPKLTLMGSNGHBFVLLKGHEDLRQDERVMQLFGLVNTLL 1469

Db   2048 --TLLKFEDEVDIMNSLQKPRKVVVRGTDGNLYPFLCKPKDDLRKDARLMEFNNLICKIL 2105

Qy   1470 ANDFTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTL-HALIRDYREKKILLNIEHRIM 1528

Db   2106 RKDQEANRRNLCLCRTYVVIPLNESCGFIEWVNHTRPFREILLKSYRQKNIPISYQEIKVD 2165

Qy   1529 LRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKL------LWL--KSPSSEVWFDRETNYT 1581

```
Db   826  --MAVCHAKLABWLVHQKKAHRTEDELQKIFHHLRRATELDKSNGSIWRTLARVHREA-  881
          |||  ||  |    || |   : :   ||: | |||::|
Qy  1062  LHYKHQNQARDEKKKLRHASGANITNATTAAITTAAITTASTEGSNSESEABSTENSPT 1121
                                     :|:|
Db   882  -------------------------ATKPADGSDSSGASG---------  896

Qy  1122  PSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDTLRVLILTWFDYGHWP---- 1177
          |:|:|              ||| :||    |     :|
Db   897  --------------------HIMEALSAYILRSVSLS--EEIEDALGFLSLWFMYG--PLLAV 934

Qy  1178  DVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLLHQLLTDIGRVHPQALIYPLT 1237
          | | |:|| |  || : |  |  || | ||:  ||||||||||||||: ||||||| |
Db   935  QVGSTIKEBIEEVNPTVWLKVVIPQIIARISSSNGTVADSVYNLLVLVVARRHPQAILYSLN  994

Qy  1238  VA-----SKSTTTARH--NAANKILKNMCE-HSN--TLVQQAMMVSEELIRVAILWEHMW 1287
          :|      | : | | :    || | || |     ||:  | | |||: | || ||
Db   995  VAHSSYQGKGTADGVEPLKGSHRVLARIAEIHQNGKAMVEDSALVCRELVRCAVLWPELW 1054

Qy  1288  HEGLEEASRLIYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRK 1347
          |||  | : |    ::: ||| |    :| |:|    | :  || |:||  | | | ||:
Db  1055  FNELGRAILYQWERQRSAENLLLAWGPLLEQLKR-PETMAEAQFAAELRQPLENACCHVER 1113

Qy  1348  YMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQVYSPKLLM-CRDLELAVPGTY--- 1404
          :| :|| |  :::| ::||:|||:|:|:|   | ||: |:|| |  |||  | ||:
Db  1114  AVSSRHEQFMEEARRIFVSIERRIREQLSGMSSLALIQLVSPKLHQNGRNLSLVVPGGYRE 1173

Qy  1405  DPNQPIRIQSIAPSLQVTTSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGL 1464
          |  |  ::    |  | |  |  || || |:|||| ||:|||||||| :|||:|||:|||
Db  1174  DGNYPL--IASFQNVLKVLNSKQRPRRIYINGSNGEIYKFLLKGHEDLRLDERVMQLLGF 1231

Qy  1465  VNTLLANDPTSLRKNLSIQRYAVIPLSTNSCLIGWVPHCDTLHALIRDYREKKKILLNIE 1524
          |||| |:||||:|:|| ||| ||||| ||  | || | | ||||||||  | :||::|
Db  1232  VNTLLEKHSVPKRHDCLIQTYSVTPLSDNAGLVGWVDHCDTLNKIIEDYRVNPR-CIRME 1290

Qy  1525  HRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSL 1584
          | : |  ||| || | |   | |:   |   | : |:  : |||  ||  | | ||
Db  1291  LDLMRSMCDNLYYLTAIQRVEPFEFPLERTEGVDLVRSFWVKAPSAETWLERRTTVVCSL 1350

Qy  1585  AVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPRLRTMLI 1644
          |:||||| ||||||||||||:|||| |:|||||||||| | | |||||| |||||:| |
Db  1351  ATMSMVGHILGLGDRHPSNLMIHAFSGRVVHIDFGDCFEVAQQRSIHPEKVPFRLTRMLV 1410

Qy  1645  NAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLN-WRLMDTNTKGNKRSR 1703
          :| | || || |  |   : |  |:| |   | |      :|    ||: : | | |::
Db  1411  KAMEMGGIEGLFRHGCHTVMNVLREEGGSLALLEAFVHDPLVSWWRDEAEGFSGNGQAA 1470

Qy  1704  TRTDSYSA-GQSVEILD-----GVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKA 1757
          :||  |:| |  |: |       || |   |   ::      |  || ||  | | ||
Db  1471  SSVQTISVIGSMTNVAEEGSVGSLQLSQLSVQRRATT----QQRVNANIEKDTSQTRKP 1525

Qy  1758  IQINRVRDKLTGRDF-----SHDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
          :|:  :   | |: |      ||  | ||: || | || |||||  | ||| |||
Db  1526  KSVVKRIKRKLRGLEFPQSQEGKSSDGFTVEEQVSRLIEEATSNENLCVHFLGWCPFW 1583
```

RESULT 10

T39911
rad3 checkpoint protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text_change 01-Dec-2000
C;Accession: T39911; T45216; S25834
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Oliver, K.; Harris, D.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21890
A;Accession: T39911
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2386 <LYN>
A;Cross-references: EMBL.AL049558; PIDN:CAB40165.1; GSPDB:GN00067; SPDB:SPBC216.05
A;Experimental source: strain 972h-; cosmid c216
R;Bentley, N.J.; Holtzman, D.A.; Flaggs, G.; Keegan, K.S.; Demaggio, A.; Ford, J.C.; Hoe
EMBO J. 15, 6641-6651, 1996

A;Title: The Schizosaccharomyces pombe rad3 checkpoint gene.
A;Reference number: Z073360; MJID:97133293; PMID:8976690
A;Accession: T45216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-197, S',199-525,'P',527-2386 <BEN>
A;Cross-references: EMBL:Y09076; NID:g1654095; PIDN:CAA70297.1; PID:g1654096
R;Seaton, B.L.; Yucel, J.; Sunnerhagen, P.; Subramani, S.
Gene 119, 83-89, 1992
A;Title: Isolation and characterization of the Schizosaccharomyces pombe rad3 gene, invo
A;Reference number: S25834; MJID:93012978; PMID:1398093
A;Accession: S25834
A;Molecule type: DNA
A;Residues: 712-1778,'LRM' <SBA>
A;Cross-references: EMBL:X63544; NID:g5023; PIDN:CAA45106.1; PID:g5024
C;Genetics:
A;Gene: SPDB:SPBC216.05; RAD3
A;Map position: 2
C;Keywords: nucleus; phosphotransferase

Query Match   7.7%;   Score 725.5; DB 2; Length 2386;
Best Local Similarity   19.6%;   Pred. No. 3.6e-37;
Matches 410; Conservative 327; Mismatches 714; Indels 639; Gaps 84;

```
Qy   37  KALILKLKDPDPDPNPGVINNVLATIGELAQVSGLEMRKKWVDELFIIMDMLQD--SSLL  94
         ::||   :   :                         |: :      | | :
Db  619  KALIFNTIKILAVENTAILETVILSWISLSRVVEBE------ELHFVILEVISSVINSGI 672

Qy   95  AKRQVALWTLGQLVASTGVV----VEPYRKYPTLLEVLLANFLKTEQN-----------  137
         |:  |::  |  | |                :  :::   |
Db  673  FYQGIGLSAL-QQIASTRHISVWQLLSPY--WPTVSVAIVQGMGKKPNIASLFAQLMNIS 729

Qy  138  QG-----------------TRRBAIRV-LGLLGALD------PYKHKVNIGMI--DQSRDA 172
                            ||       ||  |          |:     |    ||: |
Db  730  EGDFLIRTQAYTLPFLVLTKNKALIVRIABLSQSDVATLCLTNMEKILASLLTTDHPNLE 789

Qy  173  SAVSLSEKSSQDSSDYSTSEML------------VNMGNLPLDEFYPAVSMVALM--R 217
         | |:                           : |     |  :  |   | :    |
Db  790  BSVMLLLSLATHSDFEKVDLTSLLRSDPISITVELLQLYQNDVPHKIENALRKVAMIVSQ 849

Qy  218  IFRDQSLSHHHTMV----VQAITFIFKSL-GLKCVQFLPQVMPTFLNV--IRVCDG 266
         ||  |:   | |      |  || :  |   |:  | | ||  |         || |
Db  850  VVNDEDLSNKELLYDFFNNHLGILAEFSNILNDLGKGTSINEKIKTIVGIRKMLSLCGG 909

Qy  267  AIRRFLFQQLGMLVS-FVKSHIR------------------------------PYM 291
         |:  |  | |:|    |  |::
Db  910  AVKLGLPQLLSNLQSAFQNEHLRFVAIKAWFSLLLATKEPEYSSIAGLSLVILPPLFPYL 969

Qy  292  D----BIVTLMREF-------------WVMNTSIQSTII-LLIEQIVVALGGE-FKLYLP 332
         |     ::|| |:                :       |     |  | | |    | | |
Db  970  BPQEAELVIQIFDFISSDTHKCLQGLKWAIPTSLDSACFSLKAEIFCSLQNEDFYSELQ 1029

Qy  333  QLIPHMLRVFMHDNSPGRIVSIKLLAA---------------IQLFGANLDDVLHLLLPPI 378
         |:   | |                                 |  :|  | | :  || |
Db 1030  SII----KCLTNENEPVCYLGLQKLELFFQAKVDELHDTLNLDISNEVLDQLLRCLLDCC 1085

Qy  379  VKLFDAP-----------EAPLPSRKAA--------------LETVDRLTESLDF-TDY-AS 413
         |                |      | |                |  |:   | |    | |
Db 1086  VKYASTNMQISYLAAKNLGELGELGAIDPSRAKAQHIKETVVLDNFENGEESLKFILDFMQS 1145

Qy  414  RIIHPIVRTLDQSPE---------------LRSTAMDTLSSLVP------------QLGKKYQ 449
         | :|        |                                             | :|
Db 1146  QLIPAFLVTTDTKAQGFLAYALQEFLKLGGGFKSAVINKKKGLTVVTHWMSLPDLSKR-- 1203

Qy  450  IPIPM----------VNKVLVRHRINHQRYDVLICRIVKGYTLADEEDPLIYQHRMLRSG 500
         |                :  |:| |   |  ::  :|   | |
Db 1204  VLIPFLTSKYHLTPIPKIDIRYPI--YKENVTIHTWMQLFSLKLME----YAHSQ---- 1252

Qy  501  QGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDWLE-------------WLRRL 547
         |                 |:  |:|  :      | : |
Db 1253  ------------NAEKIFGICSKVVKDQEVNIPCFLLPFIVLNVIFES 1289

Qy  548  SLE-------LLKDSSSFSLRSCWALLAQAYNPMARDLFNAA-FVSCWKSBL---NED 592
```

Mol. Biochem. Parasitol. 94, 249-264, 1998
A;Title: Conserved organization of genes in trypanosomatids.
A;Reference number: Z16580; MUID:98618771; PMID:9747975
A;Accession: T14176
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1583 <BRI>
A;Cross-references: EMBL:AF031925; NID:g3452208; PID:g3452209; PIDN:AAC32769.1
C;Genetics:
A;Gene: PI3K
C;Superfamily: yeast TOR2 protein

```
Query Match             21.6%;  Score 2036.5;  DB 2;  Length 1583;
Best Local Similarity   32.1%;  Pred. No. 4.2e-121;
Matches  539;  Conservative  297;  Mismatches  643;  Indels  199;  Gaps  45;
```

```
Qy      232 VQAITFIF--KSLGLKC-VQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIR  288
                :::|| |  ::|:|:   :|:|| :::|:|||:|||:: :::::|::| :|||:::|:||
Db        5 VQVIVGIFSLEVLSPGCLVSFYGEVISSIQKQARV-QVRKREBILVHLISLVRILKGHLR   63

Qy      289 PYMDEIVTLMREF-WVMNTSIQSTILLIEQIVALGGEFKLYLPQLt-PIMLRVFMHDN  346
            |::|::| |   | :||::|:|:|:::|:|::   ::|::|:|:|::|   :|||:|||:
Db       64 PHLKEITSTVDSFISVIDLSVLRQVLALLKELCCSLRRRFRPVMSSLLGPIIVLVEERVE  123

Qy      347 SPGRIVSIKLLAAIQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTFSL  406
            :|:|:::|| ||| |||::|::|:|:|:|:|||:|||:  :|:|| |::|::||||:|:|
Db      124 ETSEIV----LDFFSAMGSLLEDHLHTVLPVVCNIIVDTVPSRCRIVAVKTLICFTKRL  179

Qy      407 -DFTDYASRIIHPIVRTLDQSP----------ELRSTAMDTLSSIVFQLGKKYQIFIPMV-  455
             :|:||:|:||||:|| | |||          |||| |||||||:|| |||||||:|:|:
Db      180 PDLCFHASRCVHCLCRVLRESDGDGVGDEGRLGCSAMRALCTLAGSLGKNFENFVLVVL  239

Qy      456 NKVLVRHRINEHQRYDVLICRIVKG-YTLADEEBDPLIYQHRMLRSGQGDALASGPVETG-  513
            :|:|::|:||||||:|:| |||:|:   || :|:::|:| :|::||:|:::::|:||:|::
Db      240 PAVADRVGETSGBY----CRFCHDIYEAIDGKRAEPVSSNGVGKAGGGGA-PSLPFTAGT  294

Qy      514 --PMKKLHVSTINLQKAWGGAARVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAY  570
              :|::|::|:|:||:|:::::::|:|||||:||:|||||:|::|||||| ||:|||:||
Db      295 SASPLKDRADAYASLRFHLRKRDQASDRDWNHLPQLAVNLLRSSSSPSHGRALTLAELH  354

Qy      571 NPMARDLFNAAFVSCKSELNEDQQDELIRSIELALTSQDI-AEVTQTLLNLAEFMEH---  626
            :|:||:||||| |||:|||:||||:|||:|||||||:||| ||:|:||||:|||:|
Db      355 EPPAAQMLHSAFAACYADMDEHTQREVIGLLTEVLRGLRVPBEVMQELLNLSEHMERQGI  414

Qy      627 --SDKGPLRDNGIV----LIGERAAKCRAYAKALHYKELEFQK-----------KHFGELE--IQA  707
              :|||:|::|::|     |:||||||||||:|||||:|||:|           |||||:|  :||
Db      415 RLSAGGKASIKHSSNFCPFDRQVLMESSANCNLYSKALHYAEIEFLETVRBYERSILRGC  474

Qy      667 -GPTPA-------ILESLISINNKLQQPEAAAGVLEYAM-------TLDVPTQVELLIKQATSHENLCQCYIGW  707
             ||:||       |:|||:|::||:|||:|||||||:||       |:||||:|||:||||:||||||||:|||
Db      475 PKPLPVEDWQNLLIKLCEKSIYFCNLLGQRESANGIIKFIRQNLPLLTGKKVTELSQMMDA  534

Qy      708 TWYEKLHBWEDALVAYDKKMDTNKDDPRLMLGRMRCLEALGEWGQLHQQCCEKWTLVNDE  767
            ||::||  |:|||:||:|:|:|||||: | ||:||::|| |||| ||:| :::||::|::
Db      535 HLFDKLRWWSQSLQAYERRLQQEPNKVSNMVGLLRALDALGEYPRVLEMWRQFSKRVSKR  594

Qy      768 TQAKMARMAAAAWGLGQWDSMBEYTCMIPRDTHDGA---FYRAVILAHQDLFSLAQQCI  824
            |::|:|:::| ||||||||||: ||:|:|::|:||:   |||||:||:|||||:||:||
Db      595 EVSKLASMGAHAWLLRRHDMEHITSFMSDEDYTGTTALFYKATIAAARKKRFREAEKLI  654

Qy      825 DKARDLLDAEHTAMAGESYSRAYGAMVSCHMLSELREVIQYKLVPERREIIRQIWERLQ  884
            |:|::||:||:|::||||||:|||:|||:||:|||::|||:|:|:|:|:||:::|:|||:
Db      655 DMCKKRVDSKLSALVAESYDRAYDLFVGIQQLSELEELAMATSDPQSAMWRQLWERRLS  714

Qy      885 GCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDP-SR  943
            |||:|||||:|:|:||::|:|||||:|||||||||||:|||||||||||:||||||| :|
Db      715 --VMAYEGWPGTLANETLLVPFSSEIDMWLRFVVSLSRAHGCQGSVSTEVLRELLGNQSIES  772

Qy      944 QLDHPLPTVHPQVTIAYMKQMWKSARKLDAFQHMQCHFYQTMQQQAGHAIATEDQQHQEL  1003
            |::||:||||||:||||::|:|:|||::|||:|||:||||||||||:||||:|:|||:||:
Db      773 AIENGIPT--PAVAMGSFQHLYETNQRDSAIARLQLIVSKVEGSGAGHVSRERED-----  825

Qy     1004 HKLMARCPLPKLGEWQLNLQGIN--ESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAV  1051
```

```
Qy     1009 RCFLKLGEW------------QLNLQ-------------------GINESTIPK-------VLQYYSAAT  1040
            : |||||  : |||||
Db     1856 KVLLKLGEWTELKSKTSNNMQVGELSFVRQQVSPQYRTKESRTPETIAFENTINYCQQAT  1915

Qy     1041 EHDRSWYKAHIAWAVMNFEAVLHYK-HQN------QARDEKKKLRHASGANITNATTAAT  1093
                   : || |:| ::: | ::|:  |::       :|:::::|::::|:|:|::|:|:|
Db     1916 QYDPGWHKVWHKLASTHFYAVCRERPHPTTVISPPQQPQQPKKWHIPP--VTRATSPPP  1972

Qy     1094 TAATATTTASTEGSNSESEAESTENSPTPSPLQGKKVTEDLSKTLLMYTVP---------  1143
                                :| :|:   ::|:|::: |
Db     1973 PA-----------------QKSPQPAPFH-SITEPLS-VIIDYPVPPLGSLVGLP  2009

Qy     1144 -----------------------AVQGFFRSISLSRGNNLQ  1161
                                   :|::|| :::: |
Db     2010 PMPAYLSSNSSLPPQHHVSPLSNDSPSNSAENKLYLKHAHAVRCFAKALMCSPGSRLE  2069

Qy     1162 DTLRVLTLWPDYG--HWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVG-RLIH  1218
            ||||:||||||:|  :||||:||:|||| |::||||:|||:|:::|:||| |:|: |:::
Db     2070 DTLRLMQLWFDHGDDKDQDQVYPALTETIFDLPVTTWLEAIPQLMARLDCPDDQKSVQLVL  2129

Qy     1219 QLLTDIGRYHPQALIYPLTVASKSTTIARHNARANKLIKNMCEHSNTLVQQAMMVSEELIR  1278
            |:::|:|:|:|:|::||||:::|:|::|:|:|::::||:|||||::|||||:|:|||::|:|
Db     2130 RVLCEIARHRPQAVIYALTTVASRSKDVHRSKNAGTVLEKOMEYHSKLVREASLVTEELVR  2189

Qy     1279 VAILWHEMWHEGLEEASRLYFGER---NVKGMFEVLEPLHAMMERG-PQTLKETSFNQA  1333
            :|::|:||||::|:::|:||:|:|    :|||:|:|:||:::::|  ||::|::|:|:
Db     2190 CATLWHEQWHDALDLDASRVYFHRRLQDNNVQAMFDALRMNDLMQKGAPTTMKEHSFQQT  2249

Qy     1334 YGRDLMBAQEWCRKYMKSGNVKDLTQAWDLIYVHVFRRISKQLPQLTSLELQYVSPKLLMC  1393
            ||:||:|:|:|||:|:|||||||:|:|:|:|::|:::|::|:|:| |::||||||||||:|
Db     2250 YSSDLKEAGRYVQAFESSGNVKDLNQAWEIYCSVFKKLRDQLATINSLDLVYVSPNLVSA  2309

Qy     1394 RDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVPILKGHEDLR  1453
            :|:|:|:||||| :|||:|||::::||||::||||||::|:|||||:|:|:||||||||:|
Db     2310 KDLELVVPGTYDPSAPIVSIQSFSSKMNVITSKQRPRKMWIRGSNGLDYQFILKGHEDPR  2369

Qy     1454 QDERVMQLFGLVNTLLANDPTSLRKNLSLQRYAVIPLSTNSGLIGWVPHCDTLHALIRDY  1513
            ||||||||||||||||:|:|::||:|||:|||::||::|||:|||||||:||||:|:|:|
Db     2370 QDERVMQLFGLVNTLLANNSETCRRNLPIQRYSIVALSKDSGLLIGWVPNCDTLHTLVKEY  2429

Qy     1514 REKK-KILLNIEHRIMLRMAPDYDHLTLMQKVEVFHAVNNTAGDDLAKLLWLKSPSSEV  1572
            ||||  |:|:||:|:|||:|||::||:|||||:|:||||:|||:|||||||||||||:|
Db     2430 REKKAKIPLSIEHKTLQKLSLETEHJTLMQKLQLFFSALSVTQQEDLRHVLWLKSPSSEV  2489

Qy     1573 WFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFP  1632
            ||||||||||:|| |||||||||||||||||||||||:|||::|:|||||||||:||:||
Db     2490 WFDRRTNYTRSVACMSMVGYILGLGDRHPSNLMLDRLITGKVVHIDFGDCFEVAMLREKFP  2549

Qy     1633 EKIPFRLTRMLTNAMEVTGLLDGNYRITCHTVMEVLREHKDSVMAVLEAFVVDPLLNWRLM  1692
            |::|||||||:|||||||||||:||:|:|||:|:||::||:||:|||:|||||:|:||||:
Db     2550 ERVPFRLTRMLINAMEVTGLLDGVVNYTAERVLKMLRTNQESLLAVLEAFVVDPVINWRLV  2609

Qy     1693 DTNTKGNKRS-RTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVXPB  1751
            :|::|||||  |:|:|||:|||:|||:|||||||||||||||:|||:|||||||||:||
Db     2610 B----GMKKDPKTRKD--TGGR-----------------QNMAGAVLPSS  2636

Qy     1752 ALNKKAIQIINRVRDKLTGRDFSHDD------TLDVPTQVELLIKQATSHENLCQCYIGW  1805
            :::|:||:||||||||||| | |:|:      |||||||||:|||||:||||||||:|||
Db     2637 ---STTDSIMETIKRKLDGTBFVHTDGSTPBPLQVTEQLAMLTEQATSPLNLCQSYIGW  2693

Qy     1806 CPFW 1809
            ||||
Db     2694 CPFW 2697
```

RESULT 9
T14176
probable phosphatidylinositol 3 kinase - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Nov-2000
C;Accession: T14176
R;Bringaud, F.; Vedrenne, C.; Cuvillier, A.; Parry, D.; Baltz, D.; Tetaud, E.; Pays, E.;

```
Db   1875 ||| ||  ||||  |||||| :: : ||  ::: :|| ||| |||| 1934
          LIQSLLIRIGENHPQALMYPLLVACKSISNLRRAAAQEVVDKVRQHSGALVDQAQLVSHE

Qy    743 MQQSGQARLEQHSAKMIAQLAKQSPKMRPYVGSLMIAMIPKLRNDQKYAE--VTAQVLN 800
          || ::|  ||   ::: : |:| ||  :|:||| |: ::|| | |  |   :|| |||
Db   1276 LIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQ-----TLKETSFN 1331

Qy     61 TIGELAQVSGLEMRKWVDELFIIMDMLQDSLLAKRQVALWTLGLVASTGYVVEPYRK 120
          || ::|  ||  ::|:| |  |||  :|:||| |: ::|| | |  |   :|| |||
Db   1935 LIRVAILWHEMWHEALEEASRLYFGEHNIEGMLKVLEPLHDMLDEGVKKDSTTIQERAFI 1994

Qy    801 AVSEIAVIGGAEIVKNLKPLFEKLTHMINDSSSLHKREAALRAIGGICRSTAYVVDPYRD 860
          |  :          ||: | :|    | :::  |: :  ||  | |:|  |:| | |
Db   1332 Q--------AVGRDLMEAQEWCRKKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTSLE 1382

Qy    121 YPTTLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
          |       ||  |  |||:|:  | |  | |||:  |: ::|  |:: | |: :| :
Db   1995 EVGNFNPGFLAYRHELKEAHECCCNYKITGKDAELTQAWDLYHVFHVFKRIDKQLASLTTLD 2054

Qy    861 YPSLLDDLLRILKTVMSNTMRREAIKTLGILGAIDPYTHKVFTGSV-QSSTAISTAISLP 919
          | :|  | |  | | :|  :: |:|  || :|| |||:|: ::| |    :::  :||
Db   1383 LQYYSPKLLMCRDLELAVPGTYDPENQPIIRIQSIAPSLQVTTSKQRPRKLTLMGSNGHEF 1442

Qy    181 KSSQDSSDYSTSEMLI--VNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHITMVVQAITPI 238
          |:||:|:: |   : ::|:: :   ||| | | |:|: :: ||: |    ::||  | ||
Db   2055 LESVSPRLLCRDLELAVPGTYRADAPVTISSFSRQLVVTTSKQRPRKLTIHGNDGBDY 2114

Qy    920 ISETDSKD-PRQDIIHWFNYEKCTLEEFYPAITIANLMLMMQDED-SQSYAELIAQAIVTI 977
          |:|||||| || |:|| :|| :|  ||||||| || |  | ||    ||  |  ||| |:
Db    239 FKSLGLKCVQFLPQVMPTFLNVIRVCDGA---------IREFLFQQLGMLVSFVKSHIRPY 290

Qy    978 FRSLGDMAPLYTEQVIP---RLIEVCRRATESSNRANLREFFLQQLANFVAIIRKHAAPY 1034
          | ||||||||||||  |   |::|| ||  ::|  || | || ||  | | :||| | |
Db    291 MDEIVTLMREFWVMNTSIQSTIILLEQIVVALGGEFKLYLPQLIPHMLRVFMEDKSPGR 350

Qy   1035 MPAIFTTIADAWKEDISVKMVVIVEVLTDMGTAIGNDFSKYTGELIPYLLTVLQTDKTKER 1094
          | || |: |  |:  |:  : :|:::  | |   |:  ::| ||  |||  || |:||  |
Db    351 IVSIKLLAAIQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTD 410

Qy   1095 VLTVKVMESIQKLTHCIVQCHLHIVLPPLLILLDDFSLKLSIRNTALSTVLHMTQQVVDVSA 1154
          |:::|: |  ||   : :||| : :| |||  | :| :::  :|: | |:| |::: ||
Db    411 YASRII----HPIVRTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVL-----VRH 462

Qy   1155 YAPRMWQSWHHNI----STAEMRDKLLLLIEIIKQLGKFFDIFKRGVDQKLRDYNLDK 1209
          | ||| |  |::|    || : ||| ||::|| :||:| |::|:|:  ::|| ::|:|
Db    463 RINHQRY--------DVLICRIVKG----------YT.ADEEEDPL 490

Qy   1210 SVHYEQYRKLAQRAQMSRDVLTSSVFAGSNGNIQYSSTQAGMRGQANNVYANND------ 1263
          |  :| |:|| |   ||| ||:| |  | | :|||: |: :|| : |:|::
Db    491 IYQHRMLRSGGQDALAS-----------GPVETGPMKKLHVSTINLQK- 527

Qy   1264 --LHERLMNGSIDSGASRQDNRDDYYRYGVEEKKEVPKVAPTTARPTSELVTVQITKQRL 1321
          | :| ::  :|   |::: |::|||| |  |::| :| : | :|:  | |:|:   |
Db    528 ------AWGAARRVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNA 580

Qy   1322 NKDALMPQWKNENLTSKDEWLQWLMKIRIGFLTYGSSPSLRAASSLGDQHPHLARDLFPA 1381
          | ||:  | |:|  :|::| |||::|||: | |:|::|:||: :| |:  | |||:|| |
Db    581 AFVSCWSELNEDQQDELIRSIELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGI 640

Qy   1382 AFMSVWTELDSDVQNDL-TSCLLRAISTGIPELIQTIINLABFMDHSERGPLPISHD--- 1437
          || |:| ||:|||:||| :|  | |::|:| |:|  ::|| ||::: |||| | ||
Db    641 VLLGERAAKCRAYAKALHYKELEFQKGP-----------TPAILESLISINNKLQQPB 687

Qy   1438 -VLGRWAEQTYAFAKACRYKEMSVLKKSGSMQTTFRKVKLEPNDCQSLITYANKLNVQE 1496
          :|| | |:|| | :| |||:|   ||:|::| :|::::| |  |:| |: :: | ||
Db    688 AAAGVLEYAMKHFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPEIM--------- 737

Qy   1497 EAAGVVRYAERNEMNFQMRGRWYEKLNEWEKALGAYELEKKKSSCPNLQVYDEKDHLMT 1556
          ::||::|| ||::||  :||||||||||||:| ||| | |:| | |::|: |:::| |
Db    738 ------LGRMRCLEALGEWGQLHQQCCEKWT---------LVNDETQA----KMA 773

Qy   1557 PEBAATEAARMHEMRCLEALGRWDELNSKSV-VWADQRGNRNDSVRDEINKKQLDHKVA 1615
          |  |  ||||:|: ||||||| | :| :|:| :| |:|| | ::||:|||:|| | |:
Db    774 RMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALIQDLFSLAQQCIDKARDLLDA 833

Qy   1616 VIAARGAWAVDNWERMADYVSVISENTQDGAMLRAVVAVHNDENTKAMGLIEKVREMIDS 1675
          | |||  || ||||| || |||  |: | | | |:||:| |:|  |  ::|:|| | |:
Db    834 ELTAMAGESYRSRAYGAMVSCHMLSELEEVIQYKIVPERREIIRQIWWERLQCQRIVEDW 893

Qy   1676 ELTAMANESYERAYIPMVSVQQMABLEEAIEYKTRPERRPRIALLWSRRLQSCRRNVEQW 1735
          |||||| ||: | || ::: :: | ||| :|||| |:: | ||| |:|||:|  | ||
Db    894 QKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKITLVLLLGVDPSRQLDH-PLPTV 952

Qy   1736 QRLIMLRGLVLLSPQEMHPLRVVFSSMCRRQGKNSMSRAVLRELLSIPANSDLVRAKAPFD 1795
          |::::||| || |::|  ||:::||:| ::| ||:|:|||| || | |::| | || || |
Db    953 HPQVTYAYMKNMWKSARKIDAFQMHQHFVQTMQQQAQHAIATEDQQ----HKQELHKLMA 1008

Qy   1796 KPLLVLALAXQLYQDDHKDEAIRALEDLANHWNKRINPIPKATGRELIPPSTKEPARICA 1855
          :: || | :|| | :| |::|::: :::  |  || :| | |:  | |:
Db      1 LEHSGITGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA 60
```

```
Query Match          36.3%;  Score 3417;  DB 2;  Length 2697;
Best Local Similarity 38.2%;  Pred. No. 8.6e-209;
Matches  781; Conservative  336;  M:smatches  603;  Indels  324;  Gaps  45;
```

```
Db  2091 AKFEPLPSVISSKQRPRKFSIKGSDGKDYKYYLKGHEDIRQDSIVMQLFGLVNTLLKNDS 2150
              ::::::::|::|| ||: :|:|||||   ::||||||:|||:|:| :||::|||
Qy   234 AITFIFKSLGLKCVQPLPQVMPTFLNVIRVCGAIREFLFQQLGMLVSFVKSHIRPYMDE  293

Db   906 SLMIIFK------------------------VLPELFHTVRTSDENLKDFITWGLGTLVSIVRQHIRKYLPE  953
              ::|:|||:                        ||:
Qy   294 IVTLMREFWVMNT---------------------SIQSTIILLEQIVALGGEFKLYLP  332

Db  2151 ECFKRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHRDAKKIPLNIEHWVMLQMAP 2210
              ||||| :||| :|||:||| ||||||| |:||||: ||::| ||:| ||:||:|||:
Qy   954 LLSLVSELWSSFTLPGPIRPSRGLPISYGEYFASLTDNQVLHLLEHLCLAINDERRTYLP 1013

Db  1534 DYDHLTLMQKVEVFEHAVNNTAGDLIAKLLWLLKSPSSEVWFDRRTNVTRSLAVWSMVGYI 1593
              || ||:|||::|:|:| :||| ||:||||||: |  ||
Qy   333 QLIPHMLRV----------FMHDNSPQRIVSIKLLAA:QLFGANLLDYLHLLLPPIVKL  381

Db  2211 DYENLTLLQKIEVFTYALDNIKGQDLYKLLWILKSRSSETWLERRTTYVTRSLAVWSMTGYI 2270
              ||:||||  |:|:  ||::  |    :   :: :|  || :::||: :| |:| |
Qy  1014 VILPCFIQVLGDAERFNDYTYVPD---------IIHFLEVFGGTLDEHMHLLPALLIRL 1063

Db  1594 LGLGDRHPSNLMLDRLSGKILHLDFGDCFEVAMTREKFPRKIPFRLTRMLTNAMEVTGLD 1653
              |||||||||||||| ||:|| |||||||||||: |    ||: ||  ||: |:|: |
Qy   382 FDAPAPLPSRKAALETVDRLTSLDFTDYASRIIHPIVRTLD-QSPELRSTAMDTLSSL  440

Db  2271 LGLGDRHPSNLMLDRITGKVIHIDFGDCFEAAILEKYPEKVPFRLTRMLTYAMEVSGIE 2330
              ||||||||||||||::|||:|:|||||||||||:|:|||||| || |:|:||:|:|::
Qy  1064 FKV-DAPVAIRRDAIKLTRVIPCVQVGHISALVHHLKLVLDGKNDELRKDAVDALCC: 1122

Db  1654 GNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQ 1713
              | ::|:|::|::|          :|         ::|  ::| |          ::
Qy   441 VFQLGKKYQIFIPMVNKLVLVRHRINHQRYDVLICR--------IVKGYTLADEE-  486

Db  2331 GSFRITCENVMRVLRDNKESLMAILEAFALDLPLIHWGFDLPPQKLTEQT-------GI 2381
              |  :|: |:||:|  :| |:|: :||  |:|  |||||||  ::|
Qy  1123 AHALGEDFTIFIESIHKLLKHRLRHKFEERIEARWRRREPLIVA--TTATQQLSRRLFV 1180

Db  1714 SVEILDGVELGEPAHKKTGTTVPESIHSPIGDLVKPEALNKKAIQIINRVRDKLTGRDF 1773
              ::||:|:|||  :    :  |:|  ||
Qy  1181 ---BDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARVSKDDWLEW 1228

Db  2382 PLPLINPSEL---LRKGAITVEREAANM---EAEQQNETKNARAMLVLRRRITDKLTGNDI 2434
              | ||:|  |                                      ::   |::
Qy  1229 EVIRDPVIENE-----------IDPFEEGTDRNHQVNDGRLRTAGEASQRSTKZDWEEW 1228

Db  1774 SHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
              :|||||||||| :|  :|| ::|:||||||||||
Qy   544 LRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRSIEL  603

Db  2435 KRFNELDVPEQVDKLIQQATSIERICQHYIGWCPFW 2470
              |||||  |||| :|  :||: |:|:|||||||||||
Qy  1229 MRHFSIELLKESPSPAIRTCAKLAQLQPFVGRELFAAGFVSCWQAQLNESSQKLVRSLEM 1288
```

RESULT 7

G96536
hypothetical protein F2J10.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G96536
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
  Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2513 <STO>
A;Cross-references: GB:AE005173; NID:g8569097; PIDN:AAF76442.1; GSPDB:GN00141
C;Genetics:
A;Gene: F2J10.9
A;Map position: 1

Query Match   40.9%; Score 3851; DB 2; Length 2513;
Best Local Similarity 42.5%; Pred. No. 1.6e-236;
Matches 830; Conservative 332; Mismatches 515; Indels 276; Gaps 38;

```
Qy     4 SGIGIKEQSARMLGHLVSNAPRLIRPVMEPILKALILLKD-PDPDDNPGVINNVLATI  62
              :|:||:::|| |||::|  |||:||| |::|  |:|:| | |:|  ||::|:: |:|
Db   677 SADNKCRREESAXLLGCLVRNCERLILLFYVAPVQKALVARLSEGTGVNANNNIVTGVLTV 736

Qy    63 GELAQVSGLEMRKWVDELFIIMMDLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRKYP 122
              ||:| ||| :|:::||| || |:|   :||:||| ||||:||:|  ||| |: | ||
Db   737 GDLARVGGLAMRQYIPELMPLIVEALMDGAAVAKREVAVSTLGQVVQSTGYVVTPYKEYP 796

Qy   123 TLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNICGMIDQSRDASAVSLSESKS 182
                 ::|  ::|:| |: |::: |::|::|| |::    |  |:
Db   797 LLLGLLLKLKGDLVWSTRREVLKVLGIMGALDPEVHK------RNQQSLSGSHGEV 847

Qy   183 SQDSSDDYSTSEMLVNMGN.PL------DEFYPAVSMVALMRIFRDQSLSHHTMVVQ 233
                |   :|::|::|:|:    ||       ::|||:|:||:||:|:
Db   848 PRGTGD-SGQPIPSIDELPVELRPSFATSEDYTSTVAINSLMRILRDASLLSYHKRVVR 905

Qy  1216 LIHQLLTIDIGRYHFQALIYPLIVASKSTTTARHNAANKILKNMCEHSNTLVQQAMMVSEE 1275
```

```
Db   604 ALTSQDI-AEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYKEL  662
              ||  |:|  :::|||||||||::|:   |:| |::|:|:||:|:|| :|: || ||: ||
Db  1289 AFSSPNIPPEILATLNLAEFMEHDEK-PLPI---DIRLLGALAEKCRVFAKALHYKEM 1343

Qy   663 BFQKGPTP--------AILBSLISINKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLH 714
              ||:||        ::|| ::::|::|| :  :  || | :|   |    :|:: | |
Db  1344 EFE-GPRSKRMDANPVAVTEALIHINNQLHQHEAAVGLITYAQQHL-DVQLKESWYEKLQ 1401

Qy   715 EWEDALVAYDKKMDTNKDDPELM----LGRMRCLEALGEWGQLHQQCCEKWTLVNDETQA 770
              :|  ||          :|:|      ||:|| |:|  ||:|:|   || |:| :|:
Db  1402 RWDDALKAYTLKA-SQTTNPHLVLEATLGQMRCLAALARWERIANNLCKEYWSPAEPSARL 1460

Qy   771 KMARWAAAAAAWGLGQWDSMEEYTCMI----------PRDTHDGA----FYRAVLALH 813
              | :       ||| |:|  |::|:|                  ::   |||: ||:|
Db  1461 EMAPMAACAAWNMGEWDQMAEVVSRLDDGDETKLRGLASPVSSGDGSNGTFFRAVLLVR 1520

Qy   814 QDLFSLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSEEVVQYKLVP----- 869
              :|| ||:|| |:|  | : ::| ::  || | |: || | | |:::|:|:
Db  1521 RAKYDEAREYVERARKCLATELAALVLESYERAYSNMVRVQQLSELEEVIEYYTLPVGNT 1580

Qy   870 ---BRRELIRQIWWERLQGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRL  926
              :     ::  |::: | | :| || || | |:|  |:  | |     ||::|  ||| |
Db  1581 IAEERRALIRNMWTQRIQGSKRNVEVWQALLAVRALVLPPTEDVETWLKFASJCRKSGRI 1640

Qy   927 ALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAYMKNMW--KSARKIDAFQHMQHFVQT 983
              || |:|  |:|  |:|:|:||| |||:|: ||:|::  :||:|:|||::|: |:
Db  1641 SQAKSTLLKLLPFDPEVSPENMQYHGPPQVMLGYLKTQWSLGEERKRKEAFTKLQILTRE 1700

Qy   984 MQQ--QAQHAIATEDQOHKQELHKLMARCFLKLGEWQLNL-QGINESTIPKV-QYYSAAT 1040
              ::   :    :  | :|  :|:|::| || |:| ::::|: |  |:: :  | |
Db  1701 LSSVPHSQSDILASMVSSKGANVPLLARVNLKLGTWQWALSSGLNDGSIQEIRDAFDKST 1750

Qy  1041 EHDRSWYKAWEAWAVMNFEAVLHYKIQNQARDEKKKLRHASGANITNATTAATTAATATT 1100
              :  |||  |  ||   :: ||: |::: ||        |:: ::
Db  1761 CYAPKWAKAWHTWALFNTAVMSHYISRGQIASQ----------------------- 1793

Qy  1101 TASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFPRSISLSRG--- 1157
              :  |: | :: ::: | :           |  | |::| :  :|  |||:||||| :
Db  1794 ----------------------YVVSAVTGVFYSIACAANAKG 1814

Qy  1158 -NNLQDTLRVLTLWFDYGHWPDVNEALVBGVKAIQIDTWLQVIPLIARIDTPRPLVGR 1215
              :|:|| ||:| : :||| : |::|  :|:| ::|: |||::|::|: ||::| ||||
Db  1815 VDDSLQDILRLLTLWFNHGATADVQTALKTGFSHVNINTWLVVLPQIIARIHSNNRAVRE 1874
```

submitted to the Protein Sequence Database, September 1995
A;Reference number: S57085
A;Accession: S57085
A;Molecule type: DNA
A;Residues: 1-2470 <MAN>
A;Cross-references: EMBL:Z49566; NID:g1015742; PIDN:CAA89594.1; PID:g1015743; MIPS:YUR06
R;Cafferkey, R.; Young, P.R.; McLaughlin, M.M.; Bergsma, D.J.; Koltin, Y.; Sathe, G.M.;
Mol. Cell. Biol. 13, 6012-6023, 1993
A;Title: Dominant missense mutations in a novel yeast protein related to mammalian phosp
A;Reference number: A54428
A;Accession: A54428
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-57, 'G',59-114,'I',116-132,'N',134-395,'K',397-546,'S',548-1467,'R',1469-16
A;Cross-references: GB:L19540; NID:g408955; PIDN:AAB66881.1; PID:g408956
A;Note: the authors translated the codon CGG for residue 1468 as Ala
R;Helliwell, S.B.; Wagner, P.; Kunz, J.; Deuter-Reinhard, M.; Henriquez, R.; Hall, M.N.
Mol. Biol. Cell 5, 105-118, 1994
A;Title: TOR1 and TOR2 are structurally and functionally similar but not identical phosp
A;Reference number: S43940
A;Accession: S43940
A;Molecule type: DNA
A;Residues: 1-230,'R',232-395,'K',397-546,'S',548-674,'I',676-1291,'E',1293-1435,'A',143
415-2470 <HEL>
A;Cross-references: EMBL:X74857; NID:g468738; PIDN:CAA52849.1; PID:g468739
R;Huang, M.E.; Manus, V.; Chuat, J.C.; Galibert, F.
Yeast 12, 869-875, 1996
A;Title: Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open reading frames
A;Reference number: S71676; MUID:96437976; PMID:8840504
A;Accession: S71688
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2470 <HUA>
A;Cross-references: EMBL:L47993; NID:g1019675; PIDN:AAB39292.1; PID:g1019688
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:TOR1; DRR1
A;Cross-references: SGD:S0003827; MIPS:YUR066w
A;Map position: 10R
C;Function:
A;Description: required for translation initiation; required for G1 progression
C;Superfamily: yeast TOR2 protein
C;Keywords: DNA binding; EF hand; leucine zipper; phosphotransferase

Query Match   42.1%;   Score 3962;   DB 2;   Length 2470;
Best Local Similarity 44.4%;   Pred. No. 1.3e-243;
Matches 815;   Conservative 360;   Mismatches 551;   Indels 110;   Gaps 26;

Qy   1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRYMEPILKALILKLKDPDDPNPGVINNVLA 60
            |: ||||:| :|| |::|| ::: :| :|| |||| :| |: ::| | ||::||||
Db   718 LKFSTSSREKEETASLLCTLIRSSKDVAKPYIRSPLLNVLLPKFQ---DTSSTVASTALR 773

Qy   61 TIGELAQVSGLEMRKWVDELFTIIMDMLQDSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
            |||||| ||::|::| ::|| : :|: : ||| ||||:: ||||||||| |:|||
Db   774 TIGELSVVGGEDMKIYLKDLPFLIIKTFQDQSNSFKREEALKALGQLAASSGYVIDPLLD 833

Qy   121 YPTLLEVLIANFLKIEQNOGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
            || ||:|| |||: | || |::|: ||:|::||:||: :|::|: ||:: |::|
Db   834 YPELLGILVNILKTENSQNIRRQTVTLIGILGAIDPYRQK--------BREVTSTTDI 883

Qy   181 KSSQDSDSDYSTSEMLVNMGNLPL-DEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITPIF 239
            ||:||:| | : | |: :| |:: :| |::| |:||| :: :|||:|::|:|:|
Db   884 SIRQNAPPIDIA--LLMQGMSPSNDEYYTTVVIHCLLKILKDPSLSSYHTAVIQAIMHIF 941

Qy   240 KSLGLKCVQPLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR 299
            | :||| ||:||: |:| :|:|| ||:|:|| |:| |:|:||| |||| :|:||::||
Db   942 QTLGLKCVSFLDQIIPTTLDVMRTCSQSLLEFYFQQLCSLIIVRQHIRPHVDSIFQAIK 1001

Qy   300 EFWVMNTSIQSTIILLEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAA 359
            ||| || :: : ::| |: ::| | : | : :| : ::|
Db   1002 DFSSV-AKLQITLVSVIRAISKALEGEFKRLVPLFLTLFLIVILENDKSSDKVLSRRVLRL 1060

Qy   360 IQLFGANLDDYLHLLLPPIVKLFDAPEARLPSRKAALEIVDRLTESLDFTDYASRIIHPI 419
            ||||||||||||||||||||||||||||||||||||||||||
Db   1414 QSIAPSLQVITSKQPRPKLITLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDP 1473

Qy   1061 LESFGPNLEGYSHLITPKIVQMAEFTSGNL--QRSAIITIGKLAKDVDLFEMSSRIVHSL 1118
            :: ||| ||| | ::|| | ||: :: ||||
Db   420 VRTLDQ--SPELRSTAMDTLSLLVFQLGKKYQIPIPMVNKVLVRHRINHQRYDVLICRIV 477

Qy   478 KGYTLADEEEDPLIYQHRMLRSGQDALASGPVETGP--MKKLHVSTINLQKAWGAARRV 535
            :: ::| |::: ::|
Db   1179 NNDVLPT----------KILEANTTDYKPAEQMEAADAGVAKLPINQSVLKSAWNSSQQR 1228

Qy   536 SKDDWLEWIRRISLELLKDSSSPSLRSCWALIAQAYNPMARDLFNAAFVSCWSLNEDQQD 595
            :| | : :|||| :|||:||::::||:|:|
Db   1229 TKEDWQEWSKRLSIQLLKESPSHALRACSNIASMYYPIAKELPNTAFACVWTELYSQYQE 1288

Qy   596 ELIRSIELALTSQ-DIAEVTQTLLNLABFMEHSDKG-PLPLRDDNGIVLLGERAAKCRAY 653
            |||||||| ||:: ::|||:| :|: ||:: : |||::|
Db   1289 DLIGSLCIALSSPLNPPEIHQTLNLVEFMEHDKALPIPTQS------LGEYAERCHAY 1342

Qy   654 AKALHYKELEFQKGPTPAILESLISINNKLQQPEAAAGVLBYAVAMKHFGELEIQATWYEKL 713
            ||||||||||:|||||:|:|||||||||:|||||||:||:|:||||||||||:|||::|||
Db   1343 AKALHYKEIKFIKEPENSTIESLISINNQLNQTDAA!GILKHAQQHH-SLQLKETWFEKL 1401

Qy   714 HEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWQLHQQCCEKWTLVNDETQAKMA 773
            || |||||:| |:| :|:| : ::| |:|:| ::|:||
Db   1402 ERWEDALHAYNEREKAGDTSVSVTLGKMRSLHALGEWEQLSQLAARKWKVSKLQTKKLIA 1461

Qy   774 RMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDA 833
            ::| |::| || : :|: |: : : |::: :: : ||:|::::|:|:|
Db   1462 PLAAGAAWGLGEWDMLEQYISVMKPKSPDKEFFDAILYLHKNDYDNASQHILNARDLLVT 1524

Qy   834 ELTAMAGESYSRAYGAMVSCHMLSELEBVIQYKLVP--ERREIIRQIWWERLQGCQRIV 890
            ||: : | ::: ::| :: :| :|: :::| |
Db   1522 EISALINESYNRAYSVIVRTQIITEPEEIIKYKQLPPNSEKKLHYQNLWTRKLLGCQKNV 1581

Qy   891 EDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLL--GVDPSRQLDHP 948
            |: :|: ||:| ::: |::|:: : :: |::::|:
Db   1582 DLWQRVLRVRSLVIKPKQDLQIWIKFANLCRKSGRWIANKALNMLLEGGNDPS------ 1635

Qy   949 LPIVH---PQVTYAYMKNMWKSARKIDAFQEMCHFV----------QTMQQQAQHAIAT 994
            |:: :: |:| :||: ||:|:| : | |:::
Db   1636 LPNTFKAPPVVYAQLKYIWATGAYKEALNEHIGFTSRLAHDLGLDPNNMIAQSVKLSSA 1695

Qy   995 EDQQHKQELHKIMARCFLKLGEWQLNLQGINESTIP-KVLQYYSAATEHDRSWYKAWHAW 1053
            :| : ||:|::|:| ||: :::| ::| |:| :
Db   1696 STAPYVBEYTKLLARCFIKQGEWRIATQPNWRTNPDAILGSYLLATHFDRKNWYXAWHNW 1755

Qy   1054 AVNFEAVLHVKHCNQARDEKKKLRHASGANITNATTAANTAATATTTASTEGSNESESA 1113
            :|:|: |:||:| :||: : :| |
Db   1756 ALANFEVI--------SMVQEETKLN--GGKNDDDDDTAVN----NDNVRIDGSIIGSGS 1801

Qy   1114 ESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSIGSLSRGNNLQDTLRVLTLWFDY 1173
            | : ::::| : ::| |: : |: : ::|
Db   1802 LVTINGNRYPLELIQR----------HVVPAIKGFFHSISLLETSCLQDTLRLILTLLFNF 1850

Qy   1174 GHWPDVNEALVEGVKAIQIDVWLQVIPQLIARIDTRPLVGRLIHQLLTDIGRYHPQALI 1233
            | |:| :| |: |: :||: :::|::|:|: | :::|||:::||
Db   1851 GGIKEVSQAMYEGFNLMKIENWLEVLPQLISRIHQPDPTVSNSLLSLLSDLGKAHPQALV 1910

Qy   1234 YPLTVASKSTTARHNAANKLLKNMCEHSNTLVQQAMMVSEELIRVALIWHEMWHEGLEE 1293
            ||||:|:|:::: |:|||::|:|::||:|:||:
Db   1911 YPLTVAIKSESVSRQKAALSIEKIRIHSPVLVNQAELVSHELIVSHELIRVALVWHELMWYEGLED 1970

Qy   1294 AGRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGN 1353
            |: |::|::|:|||:::::: ::: :: :|: ::
Db   1971 ASRQFFVEHNIEKMFSTLEPIHKHLGNEPQTILSEVSFQKSFGRDLNDAYEWLNNYKKSKD 2030

Qy   1354 VKDLTCAWDLYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGYTYDPNQPIIRI 1413
            :: |::||:|:|:||:: ::::|
Db   2031 INNLNQAWDIYYNVFRKITRQIPQLQTLDLQHVSPQLLATHDLELAVPGYYYFPGKPTIRI 2090

```
A;Cross-references: SGD:S0001686; MIPS:YKL203c
A;Map position: 11L
C;Superfamily: yeast TOR2 protein
C;Keywords: phosphotransferase

Query Match     43.6%;     Score 4105;  DB 1;  Length 2473;
Best Local Similarity  45.5%;   Pred. No. 9.5e-253;
Matches 847;  Conservative 333;  Mismatches 513;  Indels 170;  Gaps  28;

Qy      1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALIILKDPDPDPNPGVINNVLA 60
            : ||||||  :  ||: ||: :    |  :||  |::| : ::|:|
Db    727 LKFSNMPKKKEESATLLCTLNSSDEVAKPYIDPILDVILPKCQ---DASSAVASTALK 782

Qy     61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
            ::|||:  |  ::    |: |||: :    ||:  :  |: | | |||||||||  |
Db    783 VLGELSVVGGKEMTRYLKELMPLIINNIFQDQSNSFKRDAALTTLGQLAASSGYVVGPLLD 842

Qy    121 YPTLLEVLINFLKTEONGTIRREAIRVLGLLGALDPYKGHKVNIGMIDQSRDASAVSLSES 180
            |||||  |: |:|:  ::  : | |  ||:||| | |:||
Db    843 YPELLGILINILKTENNPHIRRGTVRLIGLILGALDPYKHR---------EIEVTSNS 890

Qy    181 KSSQDSSDYSTSEMLVNMGNLPL-DEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIF 239
            |||::: |::  :||::  | |: ||| ||   ::  | | | |::||  |:   |: |
Db    891 KSSVEQNAPSIDIALLMQGVSPSNDEYYPTVIHNLMKILNDPSLSIHHTAAIQAIMHIF 950

Qy    240 KSIGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR 299
            | ||| |  ||||   |::  |: |  |:| ||:|||: | |:||   | :|: :|
Db    951 QNLGLRCVSFLDQIIPGIILWRSCPPSQLDFVFQQLGSLISIVKQHIRPHEVEKIYGVIR 1010

Qy    300 EFWVMNTSIQSTILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAA 359
            |: : ::   : : :||:|  |||| ||  | |: |  || |::|:|  |:::  :|||
Db   1011 EFFPI-IKLQITIISVIFSISKALEGEFKRFVPETLTFFLDILENDQSNKRIVPIRILKS 1069

Qy    360 IQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPI 419
            : |   ||::|:|: | |:  |:  |: : || ::: :  |   ||  :
Db   1070 LVTFGPNLEDYSHLIMPIVVRMTEYSAGSL--KKISITLGRLAKNINLSEMSSRIVQAL 1127

Qy    420 VRTLDQSP-ELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVK 478
            |: | || | ::|   |||:| ||||: :: |  |  ||| |:|  | :: |
Db   1128 VRILNNGDRLTKATMNTLSLLILQLGTDFVVFVPVINKALLRNRIQHSVTDQLVNKLLN 1187

Qy    479 GYTLA-----DEEEDPLIYCHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAA 532
            | ||      |: :   |  ::   ||:  |:    :| | :::||||::   ::|  |
Db   1188 NECLPTNIIFDKENEVP--------ERKNVEDEM---------QVTKLPVNQNILKNAWYCS 1232

Qy    533 RRVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAINPMARDLFNAAFVSCKWSELENED 592
            || : |  |: ||::|:  |: | :  |:  | |  | |:|| ||:  :| || ::|:
Db   1233 QQKTKEDWQEWIRRLSIQLLKESPSACLRSCSLSVVSVYYPLARELFNASTFSSCCWVELQTS 1292

Qy    593 QQDELIRSTELAL-TSQDIAEVTQTLLNLABPMEHSDKGPLPLRDDNGIVLLGERAAKCR 651
            ||:|||  | |: | ::||  |::: ::  || :|:| |||:  |:|| |:|||:: |:
Db   1293 YQEDLIQALCKALSSSNPPEIYQMLLNLVEFMEHDDK-PLPI---PIHTLGKYAQKCH 1347

Qy    652 AYAKALHYKELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYE 711
            |: |||||||:|   ||::|: :: |  |::  |  ||::|| || ||:|||  |  |
Db   1348 AFAKALHYKEVEFLEEPKNSTIEALISINNQLHQTSAIGILKEAQQH-NELQLKETTWYE 1406

Qy    712 KLHEWEDALVAYDKKMDTNKDDPELIMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETOAK 771
            || |: |||  | :|:  |  | ||:|  |::| ||| ||: |:|    |||  |::::|
Db   1407 KLQRREDALAAYNEKRAAGEDSVEVVMGKLRSLYALGEWEELSKIASEKWGTAKPEVKKA 1466

Qy    772 MARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHODLFSLAQQCIDKARDLL 831
             | | ||| | | |||:  : | |  | | |: : | :| :||| ::||: : :| | ::
Db   1467 MAPL-AAAAWGLEQWDEIAQYTSVMKSQSPDKEFYDAILCLHRNNPKKAEVHIFNARDLL 1525

Qy    832 DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVP---ERREIIRQIWWERLQGCCQR 888
            | | | ||: |  ||: | ::|||  |:| :| :|    ::|: : || |||:   ::|
Db   1526 VTELSALVNESYNRAYNVVVFAQIIAELEEIIKYKLPQNSDKRLTMREFWNTRLLGCCQK 1585

Qy    889 IVEDVWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDSRQLDHP 948
            ::                 | | | :| | |  |: | |   | ::|:| :  :  |:|
Db   1586 NIDWQRLRVRSLVLKPKEDAQVRIKFANLCRKSGRMALAKKVLNTLL--EETDPDEDHP 1643
```

```
Qy    949 -LPTVHPQVITYAYMKQNMWKSARKIDAFQHMQHF--------------VQTMQQQAQHA 991
            :|:||||| |   :    ::::  |::                       : : :   :
Db   1644 NTAKASPPVVTAQLKYLWATGLQDEALKQLINFTSRMAHDLGLDPNNMIAQSVPQQSKRV 1703

Qy    992 IATEDQOHKQELHKLMARCFLKLGEWQLNLQGINESTIPK--------VLQYYSAATEHD 1043
            :   : ||:: |  ::: ||  |   :|:  ::|::|||          | ||: | : |
Db   1704 -----PRHVEDYTKLLARCFLKQGEWRVCLQ-----PKWRLSNPDSILGSYLLATHFD 1751

Qy   1044 RSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKLRHASGANITNATTAATTAATATTTAS 1103
            : ||||| | |  ||  | |:  :                       |:  :|| :|||
Db   1752 NTWYKAWHNWALANFEVI------------------------------SMLTSVSKKK 1779

Qy   1104 TEGSNSESRAESTR---NSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNL 1160
            ::  : ::|::| |   :||:|||  | :|::|:::|:||:|   |  |::||:|:|:|
Db   1780 QEGSDASSVTDINEFDNGMIGVNTFDAKEVHYSSNLIHREVIPAIKGFFHSISLSESSSL 1839

Qy   1161 QDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQL 1220
            ||:||:|: || |  :||  || : | |: ||||:||::|:|  |  || ||: |:: :|
Db   1840 QDALRLLT.WFTFGGIPEATQAMHEGFNLIQIGTWLLEVLPQLISRIHQPNQIVSRSLLSL 1899

Qy   1221 LTDIGRYHPQCALIYPLITVASKSTTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVA 1280
            | |:| | |  | |:|| :|   ||:| |:  ::|:|::: |:::  ||:  | |::|:||
Db   1900 LSDLGKAHPQALIVPLMVAIKSESLSRQKAALSIIEKMRIHSPVLVDQAEIVSHELIRMA 1959

Qy   1281 ILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLME 1340
            :|::|:|: |:::|: | |  :::|||:||  : |||| | ||||:| |:| | :| :|
Db   1960 VLWHEQWYEGLDDASRQFFGEHNTEKMFAALEPLYEMLKRGPETLREISFQNSFGRDLND 2019

Qy   1341 AQEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTSLELQYYSPKLLMCRDLELAV 1400
            | ||| : :::| |:|  | ||:|  || :: |||| ::||:||  |||| | |  ||:
Db   2020 AYEWLMNYKKSKDVSNLNQAWDIYYNVFRKIGKQLPQLTELQHVSPKLLISAHDLELAV 2079

Qy   1401 PGT-YDPNQPPIRIQSIAPSLQVITSKQRPRKLT.MGSNGHEFVELLKLKGHEDLRQDERVM 1459
            ||: | | || |::: |::|||::|||||:||:  |:|||  |:| :|  |:|::|:| :|
Db   2080 PGTRASGGKPIVKISKPEPVFSVISSKQRPRKFCIKGSDGKDYKYVLKGHEDIRQDSLVM 2139

Qy   1460 QLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGGWPHCDTLHALIRDYREKKKI 1519
            ||||||:|||:: | |||::|  ||| ||| | | |::|  ||: |:| |:||:|::|||
Db   2140 QLFGLVNTLLQNDAECFRRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHREAKKI 2199

Qy   1520 LLNIEHRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDLIAKLLWLKSPSSEVWFDRRTN 1579
            |||:| |:| ||:||:|:|||||| |:||  |:|||:|: ||||:|||||: ::|  |:|:
Db   2200 PLNIEHWVNLQMAPDYDNLTLLQKVEVFTALNNNEGQDLYKVLWLKSRSSETWLERRTT 2259

Qy   1580 YTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPEKIPFRL 1639
            ||||||||: || ||||||||||||||||:||::| ||||||| ::|::|::|:| |||
Db   2260 YTRSLAVMSMTGYILGLGDRHPSNIMLDRIITGKVHIDFGDCPEAAILREKFPEKVPFRL 2319

Qy   1640 TRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGN 1699
            |||| |: |:||::|::::  |  ||:|::|::||: |||:||  ||||||:  |:|||
Db   2320 TRMLIYAMEVSGIEGGSRITCENVMKVLRDNKGSMAILEAFAPDPLINWG-FDLPTK- 2376

Qy   1700 KRSRTRTDSYSAGQSVEILDGVELGBPAHKKTGTTVP-ESIHSFIGDGLVKPEAL----- 1753
            |:  :  ::   |  | ::|  ||  || :|:|     |::|||||| | :|
Db   2377 -------------------------KIEEBTGIQLPVMNANELLSNGAITBEBVQRVEN 2410

Qy   1754 ----NKKAIQINRVRDKLT.GRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWC 1806
                :::|| |||::|||: | || | |:: |:|||::||:|||:|| |:| ||||
Db   2411 EHKNAIRNARAMLVLKRITDKLTGNDIRRFNDLDYPEQVDKLIQQATSVENLCQHYIGWC 2470

Qy   1807 PFW 1809
            |||
Db   2471 PFW 2473
```

RESULT 6

S57085

1-phosphatidylinositol 3-kinase (EC 2.7.1.137) TOR1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein J1803; protein YJR066w
C;Species: Saccharomyces cerevisiae
C;Date: 13-Sep-1995 #sequence_revision 01-Mar-1996 #text_change 18-Jun-1999
C;Accession: S57085; A54428; S43940; S71688
R;Manus, V.; Huang, M.E.; Galibert, F.

```
Qy   180  SKSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIF 239
               |||:|||:      :  .: :||:||||||||::||||:  |||   ||:|||
Db   777  DLQSHNNAHGVPQ--ISAAQYPSLENYAMVAVVTLIGILKDSSLSMEHSSVVQAVMHIC 834

Qy   240  KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREPLFQQLGMLVSFVKSHIRPYMDEIVTLMR 299
              |||.:   ||||||||:||||      ||| ||:|||||||||:||||||:||:||:
Db   835  SQMGSKSTVFLPQVVPTFLQVMQSLSASSAEFYFQQLTTLTSIIGPNIRDYVSDIFNLSK 894

Qy   300  EFWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAA 359
              |||||::|| ||||||| ||  |:|||:||| ||:    ::   |  |    :|
Db   895  VFWESTTSLLLVLIELIDAIAIALQDEFKFYLPQILSCMLKAFSLDNTSRRSVSYKVLQS 954

Qy   360  IQLFGANLDDYLHILLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPI 419
               :||:.|:::|:|.:|||:||:||||.:.::|||||| .|||  ::||:|||||||
Db   955  FVIFGSNIEBYMHLVLPVIIRSFERDTTPLGFRKSALKCIAQLFQSVNFSDHASRIIHPL 1014

Qy   420  VRTLDQS-PELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVK 478
              ||| ||      |||:||| ||  |:|||:|||:|||||||:|||:||||||: :::
Db   1015 VRMLGKSNGDLRAVIMDTLCAIVSQLGYDYSIFIPMVNKVLVSHKISHPAYELLVSRLLK 1074

Qy   479  GYTLADEEEDPLIIYQHRMLRSGQGDALASGPVETGP--MKKLHVSTINLQKAWGAARRVS 536
               |  ::|::  |      |:|:           :|:|      :| |  :::::||::
Db   1075 GEPL---PKDVVVKEFKPRPS--------TKPFSTQDEVLTKLPVDQASLKAAWESSQKLT 1124

Qy   537  KDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDE 596
               |: |:|:|| |:|| :| |::|:||::|||:|   |:| |     |   : ::||:|:
Db   1125 RDDWQDWIRRISIELLKESPSSALRSCSTLAGIYHPLARDLFNVSFLSCKDELIESNKKN 1184

Qy   597  LIRSIELALTSQDIA-EVTQTLINLAEFMEHSDKG-PLPLRDDNGIVLLGBRAAKCRAYA 654
               :|:|:|.:::: :  :   ||:||:|:     |      | ||  |::    |::|:|
Db   1185 LVKSIELAMNAPNISVEILQTLINLAEYMEREDHTLPIPIK------VISAHASKCNVYA 1238

Qy   655  KALHYKELEF----QKGPTPAILESLISINNKLQQPERAAGVLEYAMKHFGELEIQATWY 710
               ||||:|||:    ::   :  |||  :||:|:| ||| ||| ||| :|: ||  |  |
Db   1239 KALHYTELQFVQFVQETKEVSISTIRSLITINNHLQQSDAAVGMLQYTKEH-KQFSLKETWY 1297

Qy   711  EKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQA 770
              |||:|:|||||:|::::||:|||  ||||::  :  : |  |:|:||:  | || |:
Db   1298 EKLHRWDDALAAYEHRERGDSSFEINIGKLRCYYALGDWDHLSELAQKAWVTSEQEHRE 1357

Qy   771  KMARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDL 830
              |  :||||||||||||:|  ::| |  | :  |  ||:|  |:||  ::::||: |:||:
Db   1358 AIAPLAAAAAWGLGQWNLISEYVSAMDRDPQDKEFFSAISAVHLGQYNKAYGHIERHRDI 1417

Qy   831  LDAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYK--LVPERREIIRQIWWERLQGCGQ 887
              |: ::::|  :|:|:||| ||:| ||::||::::||    |:|| :|  |  |::| | |:
Db   1418 LVNDLSSIIGESYNRAYGIMVKSQMLSELEBIIDYKKKMQYENNLDSLKKTWKRLEGCCQ 1477

Qy   888  RIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDH 947
              |::|:|  |:: :||.:|:|:|:||||:|| |||:|||| |||||:  ||:  |   : |
Db   1478 KNVDVWHNTLRFRAILVLSPQDSPEMWIKLADLCRRSDRLKLSNQCLTYLMGRDPSNA--Y 1535

Qy   948  PLPT---VHPQVTYAYMKNMWKSARKIDAFQHMGHFVQTMQQQAHIATEDQQH----- 999
              || :   :|:  |  :|: |: : |  |:    |:||  ||  :: |::|::
Db   1536 PLDSLKLLNPHVVYTYLKYIWATDKNIAVSELEEF--TSVLSSKHGYKMGDSSKLVDIL 1593

Qy   1000 -----KQELHKLMARCFLKLGEWQLNLQ-GINRSTIPKVLQYYSAATEHDRSWYKAWHAW 1053
                   |:|:|::||| || || |::||   :| | ||:  :| :||| |:|||||||||:
Db   1594 ASSSVSSERSPLARCFHKLGKMKSLQDSVNQSVRDILNCYFYATLFDKSWYKAWHSW 1653

Qy   1054 AVMNFEAVLHYHQNQARDEKKKLRHASGANITNATTAATTATAATTASTEGSNSESEA 1113
              |: ||:|:|:|:                  :    :      |     ||
Db   1654 ALANFEVVGYI----------------------------------------------- 1664

Qy   1114 ESTENSPTPSPIQKKVTEDLSKITLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDY 1173
              |  | . :   :| |:     |:|:|   : :|    |:: :|  ::|:|||:||:|||||
Db   1665 EQTEHG--------VTQMYE---QYIVPAIKGFFHSSVLNQKNSLQDILRLLNLWFKF 1712

Qy   1174 GHWPDVNEALVBGVKAIQIDTWLQVIPGLQIIARIDTPRPLVGRLIHQLLTDIGRYHPQALI 1233
              |  .|:  |:||:| |:|:| ||| ||  ::|:|  :|||||||::|:|:||| ||||:||:
Db   1713 GEHSDVAAIVEGFSNVPMDTWLEVIPQLIARIHTSSSSSVRASVHQLLSDIGRVHPQALV 1772

Qy   1234 YPLTVASKSTTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEE 1293
              :|| : ::::| ||| |   | : | :||:||||||| :  :|||| |||:|||||||||||
Db   1773 YSLTVSSKSTNPQQKHSAKSIMDSMLSHSDTLVRQALLVSQELIRVAILWHELWYEGLEE 1832

Qy   1294 ASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGN 1353
              ||  ||  |: ||:|::: |:|:: :|| |:|:|:|| :  |||| |:|:| ||: ||||
Db   1833 ASQAYPSDHDISLMIDIVKPLHETLEKGPSTLSEISFAQTFGYDLRKARSYWQKPLQDGD 1892

Qy   1354 VKDLTQAWDLYHVFRRISKQLPQLTSELQYVSPKLLMCRDLELAVPGTYDPNQPIIRI 1413
              . :|  :||| || |||::||| |  :|| :||||:| ||  ||||| | | :|:|||:
Db   1893 PTELNQSWDLYIYQVFRRIQKLPRIKHLELQYVSPKLLDACDLELAVPGTYGHNKPVIRI 1952

Qy   1414 QSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDP 1473
              .::   ||:::|:||||: |||:|||: :||:|:|||: ||||||| ||:|| ||| |
Db   1953 SHFHHTFEVISSKQRPRRLTIHGSDGKDYQVVLKGHEDLRQDERVMQLFGLCNTLLTTDS 2012

Qy   1474 TSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYREKKKILLNIEHRIMLRMAP 1533
              | :||| |::|  || |::||||||||| | || :| |::|:|:||||| |:|| :||||
Db   2013 ETFKKRLNIERYVIPLSPNSGLLGWVPHSDTLHFLIKEFRSKRNILLNLEHRMMLQMAP 2072

Qy   1534 DYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVNSMVGYI 1593
              | | |:|::| || |:||  ::||:||| |:| | || ||  :||:| || ||:|:|||||
Db   2073 DCDSLTLLQKLEVFEYVMANTDGYDLHVLWLKSRSSEAWLDRRTSYTQSLAVMSMVGYI 2132

Qy   1594 LGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLD 1653
              |||||||||||||||||||||||||||||||||::|:||||||||||||||:||||:|||
Db   2133 LGLGDRHPSNLMDRYSGKIIHIDFGDCFEVAMHREKFPEKIPFRLTRMLINAMEVSGIQ 2192

Qy   1654 GNYRITCHTVMEVLREHKDSVMAVLRAFVDPLLNWRLMDTNVTKGNKRSRTRTDSYSAGQ 1713
              | |:||::| ||||| :::::|||||||||||:||:|:|:  |:| |:|  || ||    :
Db   2193 GTYKITCELVMRVLRSNTESLMAVLRAFVDPLINWRLMTKSSFGASTLRPTSS------- 2247

Qy   1714 SVEILDGVGLGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDF 1773
              |||   | || |  |     :                    :
Db   2248 SVE-------EKGRSYTHRARHADYAALSETNGVNAEGLNERSIQVLKRVSNKLITGKDF 2299

Qy   1774 SHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
              .| :| :||::|:|:|:|::: ||| :| ||| ||:|
Db   2300 DLKEQLPVKAQVEKLIQQATAPENLCRCYVGWCSFW 2335
```

RESULT 5
S38040
1-phosphatidylinositol 3-kinase (EC 2.7.1.137) TOR2 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YKL203c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 12-Nov-1999
C;Accession: S38040; S38041; S35106
R;Maia e Silva, A.; Bossier, P.; Vilela, C.; Fernandes, L.; Soares, H.; Guerreiro, P.;]
  submitted to the Protein Sequence Database, March 1994
A;Reference number: S38040
A;Accession: S38040
A;Molecule type: DNA
A;Residues: 1-2360 <MAI>
A;Cross-references: EMBL:Z28203; GSPDB:GN00011; MIPS:YKL203c
A;Experimental source: strain S288C
R.Pohl, T.M.; Pohl, F.M.
  submitted to the Protein Sequence Database, March 1994
A;Reference number: S37897
A;Accession: S38041
A;Molecule type: DNA
A;Residues: 1902-2473 <POH>
A;Cross-references: EMBL:Z28203; GSPDB:GN00011; MIPS:YKL203c
R.Kunz, J.; Henriquez, R.; Schneider, U.; Deuter-Reinhard, M.; Movva, N.R.; Hall, M.N.
A;Title: Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase
  Cell 73, 585-596, 1993
A;Reference number: S35106; MUID:93258821; PMID:8387896
A;Accession: S35106
A;Molecule type: DNA
A;Residues: 1-1472, 'G', 1473-2473 <KUN>
A;Cross-references: EMBL:X71416; NID:g298027, PID:g298028
C;Genetics:
A;Gene: SGD:TOR2; MIPS:YKL203c

```
Qy  1284 HEMWHEGLEPASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE 1343
         ||||||||| ||||||||:|:| |||| ||||:| |:|||||||:|| |||| |||:|| |
Db  1826 HEQWHEGLEEASRLYFGDHNIRGMFAVLRPLHHEMLERGPETLREISFQQAFGRDLVEARD 1885

Qy  1344 WCRKYMKSGNVKDLTQAWDLYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGT 1403
         :|| ::|:|:: || :|||||:::|: ||:|||||:|||||||||||: |||||||||
Db  1886 CCIRFEQTGDISDLNQAWDLYYQVFKKIRKQLPQLTTLDLQYVSPKLLHVHDLELAVPGT 1945

Qy  1404 YDPNQPIIRIOSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFG 1463
         | |.||:||| . :. :|::|||||||:|::.|| ||:| |:|||||:|||||||||||
Db  1946 YVSGKPVIRIVKFYPTFNVITSKQRPRRLSIKGSDGKDYQYVLKGHEDIRQDERVMQLFG 2005

Qy  1464 LVNTLLANDTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNI 1523
         | | ||| || ::||||||||.|||||.:||::|||| ||||:|||||||||:: |||||
Db  2006 LCNLLLLADPETFKRLLSIQRYPVIPLSPDSGLLGWVLDSDTLHVLIRDYRESRKILLNI 2065

Qy  1524 EHRIMLRMAPDYDHLTLMQKCVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRS 1583
         ||| ::||||||| |||:||.:|||| |: .|||||| :||||||:|||.|::||||| |
Db  2066 EHRLIIQMAPDYDRLTLLQKVEVFEYALLSTTGQDLYRVLWLKSRSSEAWLNRRTNYSRT 2125

Qy  1584 LAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPPEKIPFRLTRML 1643
         ||||||||||||||||||||||||| :|:|:|||||||||||| |||:|||||||||||
Db  2126 LAVMSMVGYILGLGDRHPSNLMLDRYTGNIIHIDFGDCFEVAMHREKFPPEKIPFRLTRML 2185

Qy  1644 TNAMEVTGLDGNYRITCHTVMEVI.REHKDSVMAVLEAFVYDPLINWRLMDTNTKGNKRSR 1703
          |||||:|::| . |||  :|.|: | :|:|||||||||||||||||||:   :.
Db  2186 VNAMEVSGIEGTFRITCEHVMRVLRTNKESVMAVLEAFVYDPLINWRLAPAYSPSIDEKQ 2245

Qy  1704 TRTDSYSAGQSVEILDGVELGEPAHKKTVGTTVPESIHSPIFGDGLVKPEALNKKAIQINR 1763
         ::  | || : : |: :  |   | ||  |:::||||| | || :    : :||:::|:||
Db  2246 SNEPNTLLGETIDGLHRKRLNE----EGITLEER------QKPBILNQRAITVLNR 2291

Qy  1764 VRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWCPFW 1809
         | ||||||||   :: |||:||| :|:||||:|||| |||||| ||
Db  2292 VSNKLTGRDFKPQQQLDVPSQVERKLILQATSIENLCLCYIGWCSFW 2337
```

RESULT 4
T40186
probable phosphatidylinositol 3-kinase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T40186
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21910
A;Accession: T40186
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2335 <WOO>
A;Cross-references: EMBL:Z97992; PIDN:CAB10805.1; GSPDB:GN00067; SPDB:SPBC30D10.10c
A;Experimental source: strain 972h-; cosmid c30D10
C;Genetics:
A;Gene: SPDB:SPBC30D10.10c
A;Map position: 2
C;Superfamily: yeast TOR2 protein

```
Query Match     44.1%;  Score 4147;  DB 2;  Length 2335;
Best Local Similarity   45.9%;  Pred. No. 1.8e-255;
Matches   843;  Conservative 337;  Mismatches 516;  Indels 140;  Gaps 24;

Qy     1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKDPDPDNPGVINNVLA 60
         :| || . ::|::|::| || || ||:| || .:|: :|||: ||   ||:.:::.
Db   613 MEYSGNSRQKEESAQLLKLLVSXARTLIKPYIQSIIHVILPKA---ADTSPGVSSAIIS 668

Qy    61 TIGRLAQVSGLEMRKWVDELFI-IIMDMLQDSLLAKRQVALWTLGQLVASTGYVVEPYR 119
          :| | .||| |::  | | |  :|:::||||| ::|:::: || |||||:|||::||
Db   669 ALGELASVEGEDMPVDVRGSFMKLILVNLQDQSSTLKRLASLKCLRKLCGRSGYVIQPYL 728

Qy   120 KYPTLEVLLNFLKTBQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSE 179
          || |  |: || || :|:  .||:.:|| ||||||||   :  :||:||:||:|:|
Db   729 DYPPLLGALIGILQSBQPTPIRREVLRTLGVLGALDPYTY-------LTTEVSD 776
```

```
Qy   228 HTMVVQAITFIPKSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHI 287
         ||:|:||:::|.|::|| |  ||.|::| ::.|:| | . | ||:||||.::|.|:|.||
Db   824 HTAVIQAVMYIFKTMGLRCAPFLSQIIPEFIAVMRTCPTNILEFYFQQLSIIVLIVRQHI 883

Qy   288 RPYMDEIVTLMREFWVMNTSIQSTILLIIEQIVVALGGEFKLYIPQLIPHMLRVFMHDNS 347
         | ::|::. |:::||  ::.:|.|||:::| :: || |||| ::| ::.::|..|:.| |
Db   884 RSFLPDLFKLIKDFWNPHSNLQFTILSLIESLARAMQGEFKYPLPSLLVMMLQIFDSDVS 943

Qy   348 PGRIVSIKLLAAIQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLD 407
          . |::::|:||:::||..| ||::.|||||:| ::.:   : : | | |:::::: :::
Db   944 VDSVSTKKVLHAFIVFGDTLADYFEMLLDPILRLYERNDVSGIGIKESIMTIGRLSMVIN 1003

Qy   408 FTDYASRIIHPIVRTLD-QSPELRSTAMDTLSSLVFQLGKKYQIFPIPMVNKVLVRHRINH 466
         ::.|||||||||.|:|. :: :|:|:.|||.:.|::||...: ||||||:.|.|    :.|
Db  1004 LSEYASRIIHPVMRMLSCNNASLIRVSMDTVCALIYQLNVDFAIFPIPMIDKCLKMNGVTH 1063

Qy   467 QRYDVLICRIVKGYTLADEEEDPL---IYQHRMLRSGCGDALASGPVETGPMKKLHVSTI 523
          .| .|. :.  :: ||.: :.:    ::| : | : :  ||:: |::::.:||| |:
Db  1064 ETYSILVEQFL-------QEQLPLIKLNPYEKYDKPKLDVVASAADITS--KKLPVNQE 1113

Qy   524 NLQKAWGAARRVSKDDWLEWLRRISLELLKQSSSPSLRSCWALAQAYNPMARDLFNAAFV 583
         .| :||.|:: ||||||::|:||:. ||| :|.| :||:||||| ||:|:|||||||:||
Db  1114 ILRNAWEASQRSTKDDWQEWIRRLGVALLRESPSHALRACAALAAAYQPLARDLFNASFV 1173

Qy   584 SCWSELNEDQQDELIRSIELALTSQDIA-EVTCTLLNLABFMEHSDKGPLPLRDNGIVL 642
         |||||:: .  :|::|||::||||::|: ||: :||||||||||:|| |||:    |::
Db  1174 SCWSELYDHFQEELVKSIEIALTSPHISPELIQILNLABFMEHDDK-PLPI----DIRT 1228

Qy   643 LGBRAAKCRAYAKALHYKELEFQK-----GPTPAILELSLISINNKLQQPEAAAGVLBYAM 697
         || :|||| | ||||||||||::     |:| | |:||.|||||:|||| ||| |:|:|
Db  1229 LGAYAAKCHAPAKALHYKELEFIEEELVTKPSVDTIRALISINNQLQQPDAAIGILKHAQ 1288

Qy   698 KHFGELEIQATWYEKLHEWEDALVAYDKKMDINKDDPELMLGRMRCLEALGEWGQLHQQC 757
         :|:.::  : ||| |::|||||:|| :: . ..  |.| ::|::|||.|||||::|.:
Db  1289 QH-DKMNLKETWYKPKLQRWEDALSAYEKREAAGAGNFEITMGKLRCLHALGEWDRLSQLA 1347

Qy   758 CEKWTLVNDETQAKMARMARMAAAAAWGLGQWDSMEBYTCMIPRDTHDGAFYRAVLALHQDLF 817
          |:|   .|. :|::| :. .||||||||||||:.| :|: :..:| | || :||:|||..|
Db  1348 QENWIHAGHDARRYIAPLSVAAAWGLGQWEQEQMDEYISVMKSESPDKAPFNAIVALHRSQF 1407

Qy   818 SLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLV--PERREI 874
         .::. . .:|||||| |:|::.||| .||| |:|  :|||||::|||| ||    |::|
Db  1408 EEAASYITRARDLLDTELITALVGESINRAYSVAVRVQMLSELEBIITYKKAEDKPEVREM 1467

Qy   875 IRQIWWERLQGCQRIVSDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIV 934
         |::.| :||.|||: |.||:::|::|||::||::|:|.|:| |.||.|||| :||:|::
Db  1468 IKKTWVRRLLKGCQRNVDVWQRMLRIRSLVISPRDNMEMWIKFANLCRKSGRISLAKKSLN 1527

Qy   935 LLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQHMQHFVQTMQQ-------- 986
         |||..:.| ||:.: |::||||.||:||:||||||||||||||||||||:||
Db  1528 LLLEDDENLDNSLVLKKTHPSIVYANLKFLWLAVDDKRKALNSMQEFTSQLISDINVDPAL 1587

Qy   987 --QAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQ-GINESTIPKVLQYYSAAATEHD 1043
           ||: .:..::.::.:|: :::|||::|:|||| ||.  ::| ||||::||||:|..::
Db  1588 FVQSTSVNTQKSQEEIQYYFHLLARCYHKQGQWQQEFIENNWSEGSFDGVLQSIMYATQFD 1647

Qy  1044 RSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKLRHASGANTINATTAATATTTAS 1103
         .|||||||:|| :||||| :
Db  1648 SKWYKAWHSWALANFEAVKFLE-------- 1669

Qy  1104 TEGSNSBESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSISLSRGNNLQDT 1163
                   :|:::..| |     |:|:||:|: :||||:.|||:|||:||:|:| :|||||
Db  1670 ----------QSEEQIPSAAYEQ----------YIIPAVKGFFKSIALSKG-NLQDT 1705

Qy  1164 LRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPLIARIDTPRPLVGRLIHQLLTD 1223
         || |.|||.:|: |.|.|::..|:: |:||:|| ||| |||| :.::|:|:::||||:|
Db  1706 LRLLNLWFKFGNNSNVINTLNVGISTVNIDIWLDVIPQIARIHASSLSVRKSVHQLLSD 1765

Qy  1224 IGRYHPQALIYPLTVASKSTTTARHNAANKILKXMCEHSNTLVQQAMMVSEELIRVAILW 1283
         :||:||||::|||||| || .:||:||| | |  |. | :::|:||:: |||||:||||
Db  1766 VGRAHPQALVYPLTVAAKSQSSARQNAALAIMDSLKTHSPRIVEQARLVSHELIRAAILW 1825
```

```
Db  1101 QLFGANLDDYLHLLLPPIVLFDAPEVPLPSRKAALETVDRLTESLDFTYASRIIHPIV 1160

Qy   421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy   481 TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1221 TLADEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy   541 LEWLRRLSLELLKDSSSFSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1281 LEWLRRLSLELLKDSSSFSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy   601 IELALTSQDIAEVTCTLLNLAEFMEHSDKGPLPLRDNGIVLLGERAAKCRAYAKALHYK 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1341 IELALTSQDIAEVTCTLLNLAEFMEHSDKGPLPLRDNGIVLLGERAAKCRAYAKALHYK 1400

Qy   661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1401 ELEFQKGPTPAILESLISINNKLQQPEAASGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy   721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKWARMAAAAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKWARMAAAAA 1520

Qy   781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy   841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERRBIIRQIWWERLQGCQRIVEDWQKILMVR 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 1640

Qy   901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTTYAY 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTTYAY 1700

Qy   961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy  1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy  1081 SGANITNATTAATTAATATTTPASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
         |||||||| || |:||::|:||:|||||||||||||: |::|||||||||||||||||:|
Db  1821 SGANITNATTTATTAASAAAATSTEGSNSESEAESNESSSPTPSPLQKKVTEDLSKTLLLY 1880

Qy  1141 TVPAVGGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
         |||||| |||||||||||||||||||||||||||||||||||||||||||
Db  1881 TVPAVGQFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy  1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 2000

Qy  1261 HSNTLVQQAMMVSEELIRVALIWHEMWHEWGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
         ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  2001 HSNTLVQQAMMVSEELIRVALIWHEMWHEWGSLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy  1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLIYHVFRRISKQLPQLTS 1380
         ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2061 GPQTLKETSNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLIYHVFRRISKQLPQLTS 2120

Qy  1381 LELQYVSPKLLMCRDLELAVPGTYDDNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2121 LELQYVSPKLLMCRDLELAVPGTYDDNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy  1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240
```

```
Qy  1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy  1561 KLIMLKSPSSEVWFDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMLDRLSGKILHIDFGD 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2301 KLIMLKSPSSEVWFDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy  1621 CFEVAMTREKPPEKIPFELTRMLTNAMEVTGLDGNYRITCHTMEVLREHKDSVMAVLEA 1680
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  2361 CFEVAMTREKPPEKIPFELTRMLTNAMEVTGLDGNYRTTCHTMEVLREHKDSVMAVLEA 2420

Qy  1681 FVVDPLLNWRLMDTNKTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
         |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  2421 FVVDPLLNWRLMDTNAKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy  1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy  1801 CYIGWCPFW 1809
         |||||||||
Db  2541 CYIGWCPFW 2549
```

RESULT 3
T40577
probable phosphatidylinositol 3-kinase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 15-Sep-2000 #text_change 06-Oct-2000
C;Accession: T39913; T40577
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Oliver, K.; Harris, D.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21890
A;Accession: T39913
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 365-2337 <LIN>
A;Cross-references: EMBL:AL049558; PIDN:CAB40167.1; GSPDB:GN00067; SPDB:SPBC216.07c
R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A;Reference number: Z21938
A;Accession: T40577
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-398 <SEE>
A;Cross-references: EMBL:AL035216; PIDN:CAA22805.1; GSPDB:GN00067; SPDB:SPBC646.01c
A;Experimental source: strain 972h-; cosmid c646
C;Genetics:
A;Gene: SPBC646.01c
A;Map position: 2
C;Superfamily: yeast TOR2 protein

Query Match   46.7%; Score 4398; DB 2; Length 2337;
Best Local Similarity 48.1%; Pred. No. 1.8e-271;
Matches 888; Conservative 324; Mismatches 476; Indels 158; Gaps 20;

```
Qy     1 LEHSGIRIKEQSARMLGHLVSNAPRLIRPYMEPIIKALILKLKDPDDNPGVINNVLA 60
         |::|:|:||| ::||:.|| ...:|:  || :::: :|:: :|:
Db   613 LDYSTIIRTKENAKLLCLLIAAAPRLIESHVEPILQILLPKAKDSSSIVAASIVN---- 668

Qy    61 TIGELAQVSGLEMRKWVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
         ::|:.|:|| :.|:..|| : :|:: :|  :|:::|:
Db   669 SLGEICQISGEVIVPFIKDLMPLIIEALQDQSSPIRRAAALKALGNLSSSTGYVIDPYIE 728

Qy   121 YPTLLEVLLNFLKTEQNQGITRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
         :.::.:|:|: .:.||:..|:: ::|:
Db   729 FPSLLDILIGITKEQDITIRRETIKLIGTLGALLDPNRHRV--------------- 769

Qy   181 KSSQDSSDYSTSEMLVNMGNLP---------------LDEFYPAVSMVALMRIFRDQSLSHH 227
          .::| .:. ::. ::.|::||               |||:|.|.||.|:|: |||:|:..
Db   770 -----LEKGTEKVVPEQKNIPPDISLLMSGIGPSSDEYYPPTVVTIALMSILKDPSLITIH 823
```

Qy 1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVPEHAVNNTAGDDLA 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVPEHAVNNTAGDDLA 2300

Qy 1561 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2301 KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSCKILHIDFGD 2360

Qy 1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHDSVMAVLEA 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2361 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHDSVMAVLEA 2420

Qy 1681 FVYDPLLNWRLMDTNTKGNKRSRTPTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2421 FVYDPLLNWRLMDTNTKGNKRSRTPTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH 2480

Qy 1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy 1801 CYIGWCPFW 1809
         |||||||||
Db 2541 CYIGWCPFW 2549

RESULT 2
A54837
rapamycin/FKBP12 target 1 - rat
N;Alternate names: RAFT
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 24-Nov-1999
C;Accession: A54837
R;Sabatini, D.M.; Erdjument-Bromage, H.; Lui, M.; Tempst, P.; Snyder, S.H.
Cell 78, 35-43, 1994
A;Title: RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashi
A;Reference number: A54837; MUID:94306515; PMID:7518356
A;Accession: A54837
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2549 <SAB>
A;Cross-references: GB:U11681; NID:g511228; PIDN:AAA20091.1; PID:g511229
C;Superfamily: yeast TOR2 protein

Query Match   99.2%; Score 9341; DB 2; Length 2549;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1793; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy   1 LEHSSGIRKEQSARMLGHLVSNAPRLIRPYMEPLLKALILKLKDPDPDNPGVINNVLA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 741 LEHSSGIRKEQSARMLGHLVSNAPRLIRPYMEPLLKALILKLKDPDPDNPGVINNVLA 800

Qy   61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVVPEPYRK 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 801 TIGELAQVSGLEMRKWVDELFVIIMDMLQDSLLLAKRQVALWTLGQIVASTGVVASTGVVVPEPYRK 860

Qy 121 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 861 YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy 181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy 241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy 301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1041 FWVMNTSIQST-IILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSQGRIVSIKLLAAI 1100

Qy 361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLITESLDFTDYASRIIHPIV 420

Qy 421 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1161 RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY 1220

Qy 481 TLADEEEDPLIYQHRMLRSGGQDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1221 TLADEEEDPLIYQHRMLRSGGQDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy 541 LEWLRRISLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1281 LEWLRRISLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy 601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy 661 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1401 ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy 721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1461 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy 781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy 841 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCGRIVEDWQKILMVR 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1581 ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCGRIVEDWQKILMVR 1640

Qy 901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLFTVHPQVTYAY 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLFTVHPQVTYAY 1700

Qy 961 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1701 MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy 1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy 1081 SGANITNATTNATTAATTNATTAATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1821 SGANITNATTNATTAATTTA?ATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880

Qy 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKA-QIDTWLQVIP 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy 1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE 2000

Qy 1261 HSNTLVQQAMMVSEELIRVALLWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2001 HSNTLVQQAMMVSEELIRVALLWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy 1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLIQAWDLYYHVFRRISKQLPQLTS 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLIQAWDLYYHVFRRISKQLPQLTS 2120

Qy 1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH 2180

Qy 1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2004, 19:59:08 ; Search time 28 Seconds
                                        (without alignments)
                                        6214.660 Million cell updates/sec

Title: US-09-517-491-12
Perfect score: 9413
Sequence: 1 LEHSGIGRIKEQSARMLGHI.........KQATSHENLCQCYIGWCPFW 1809

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR 78:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9413	100.0	2549	2	S45340	FKBP-rapamycin-ass
2	9341	99.2	2549	2	A54837	rapamycin/FKBP12 t
3	4398	46.7	2337	2	T40577	probable phosphati
4	4147	44.7	2335	2	T40186	probable phosphati
5	4105	43.6	2473	1	S38040	1-phosphatidylinos
6	3962	42.1	2470	2	S57085	1-phosphatidylinos
7	3851	40.9	2513	2	G96536	hypothetical prote
8	3417	36.3	2697	2	T25444	hypothetical prote
9	2036.5	21.6	1583	2	T14176	probable phosphati
10	725.5	7.7	2386	2	T39911	rad3 checkpoint pr
11	663.5	7.0	2354	2	T13288	mei-41 protein - f
12	645	6.9	1663	2	T28923	hypothetical prote
13	641.5	6.8	3856	2	T51174	ataxia-telangiecta
14	628	6.7	2368	2	S46005	ataxia telangiecta
15	613.5	6.5	1708	1	A43100	ESR1 protein - yea
16	585.5	6.2	2812	2	T43271	phosphotidylinosit
17	552.5	5.9	4128	2	JC6306	protein kinase (EC
18	535	5.7	4096	2	A57099	DNA-activated prot
19	494	5.2	2787	2	S45416	TEL1 protein - yea
20	453	4.8	930	2	G02083	DNA-PK - human (fr
21	390.5	4.1	3655	2	T38084	TRAP-like protein
22	386.5	4.1	2514	2	T37320	ataxia telangiecta
23	370	3.9	3744	2	S46715	hypothetical prote
24	356	3.8	2619	2	T24588	hypothetical prote
25	310.5	3.3	2089	2	C85426	ATM-like protein	
26	310.5	3.3	3738	2	T05501	hypothetical prote
27	304.5	3.2	3795	2	T00831	hypothetical prote
28	258	2.7	887	1	S57219	1-phosphatidylinos
29	228.5	2.4	897	2	T43628	phosphatidylinosit
30	227.5	2.4	873	2	T25442	hypothetical prote
31	226	2.4	875	1	A36369	1-phosphatidylinos
32	219	2.3	801	1	T52538	1-phosphatidylinos
33	216	2.3	664	2	PC4002	phosphatidylinosit
34	216	2.3	1043	2	T43502	1-phosphatidylinos
35	212.5	2.3	812	2	T07745	phosphatidylinosit
36	212.5	2.3	1876	2	T13801	phosphoinositide 3
37	211.5	2.2	1020	2	T18260	1-phosphatidylinos
38	205.5	2.2	814	2	T07761	phosphatidylinosit
39	205	2.2	814	2	B96630	Phosphatidylinosit
40	204.5	2.2	1093	2	T18275	1-phosphatidylinos
41	200	2.1	3944	2	T19997	hypothetical prote
42	199.5	2.1	732	2	T08420	1-phosphatidylinos
43	199.5	2.1	2121	2	T27406	hypothetical prote
44	196.5	2.1	1068	1	A43322	1-phosphatidylinos
45	196.5	2.1	1068	1	I38110	1-phosphatidylinos

ALIGNMENTS

RESULT 1
S45340
FKBP-rapamycin-associated protein (FRAP) - human
C;Species: Homo sapiens (man)
C;Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C;Accession: S45340
R;Brown, E.J.; Albers, M.W.; Shin, T.B.; Ichikawa, K.; Keith, C.T.; Lane, W.S.; Schreibe
Nature 369, 756-758, 1994
A;Title: A mammalian protein targeted by G1-arresting rapamycin-receptor complex.
A;Reference number: S45340; MUID:94277209; PMID:8008069
A;Accession: S45340
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2549 <BRO>
A;Cross-references: GB:L34075; NID:g508481; PIDN:AAA58486.1; PID:g508482
C;Genetics:
A;Gene: GDB:FRAP1; FRAP; RAFT1
A;Cross-references: GDB:597698; OMIM:601231
A;Map position: 1p36-1p36
C;Superfamily: yeast TOR2 protein

Query Match   100.0%; Score 9413; DB 2; Length 2549;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKDPDDPNPGVINNVLA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKDPDDPNPGVINNVLA 800

Qy   61 TIGELAQVSGLEMRKKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801 TIGELAQVSGLEMRKKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 860

Qy   121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy   181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK 980

Qy   241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy   301 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPLIPHMLRVFMHDNSPGRIVSIKLLAAI 1100

Qy   361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

```
Qy   636 DNGIVLLGERAAKCRAYAKALHYKELEFQKGPTALLBSLIS-----------------678
Db  2441 IGSIISFL--------SREYHIRLL----GKTPNVLEFILTSIVSSDMPIPLPPHLLV 2486

Qy   679 -----------INNKLQQP--------EAAAGVL----EYAMKH----- 699
Db  2487 YLSKTYGLIHYCILLLENSLQNNPGLSEDELIVYHKSCLDALSDIYYSLDEHDLYHGLWR 2546

Qy   700 ----FGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMCLEAL-------G 748
Db  2547 RRANFLETEV-ATSHEQCHEWEKAQLVY--------EHAQLKV-------CTGSLFYSPTEHG 2593

Qy   749 EWGQLHQQCCEK---WTLVND-ETQAKMARMAAAAWGLGQW----DSMEEYT------C 794
Db  2594 FWLDHWILCAQKLNQWDVLFDFSKQEGCAELVLEECAWRLLSDWSTEQDTLEKAIKSLSPFT 2653

Qy   795 MIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDA--ELTAMAGESY--------- 843
Db  2654 SLRRHTADALLY-----------LNKTQRKMGSVTEFSRIIDECMQFSLRRWQQ 2696

Qy   844 --SRAYGAMVSCHMLSELEEVIQ--------YKLVPERR----------EIIRQIWWERL 883
Db  2697 LPKRVYQSHVS--LLHHFQEIVELQEAFGIYSQLNDTNIHNIDNKLRDIKVVLQGWRERL 2754

Qy   884 QGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSR 943
Db  2755 PNVWDDIDIWSDLIIAWRQSVFKSIN--KVFLPIVSIAQQSTNKSNTN-------------- 2799

Qy   944 QLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQH---MQHFVQTMQQQAQHAIATEDCQHK 1000
Db  2800 -------------SVSYLY-RGYHELAWIINRFAHVARVHHLPEVCINQLTKIYTLPNIEIQ 2847

Qy  1001 QELIKL--MARC-FLKLGEWQLNLQGINESTI---------------PKV 1032
Db  2848 EAF+KLRBQAEBCHYESPSEMQLGLEVINNTNLMYFRNRQKAEFFTLKGMFQNRLGEKDEA 2907

Qy  1033 LQYYSAATEHDRSWYXAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAA 1092
Db  2908 NQAFATAVOIDIGSGXAWSEWGLYHDE-------------------LFQANPQEIHHACNAV 2950

Qy  1093 TTAATATTTASTEGGNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFPRSI 1152
Db  2951 SCFIQASSLLSS--SNSK------------PL-------ITRVLMLLSVDDSHG---- 2983

Qy  1153 SLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPL 1212
Db  2984 ------------SVSEVVSSFKSEIPTWNWIPFIPQLLSALSHRESI 3018

Qy  1213 VGRLIHQLLTDIGRYHPQALIYPLTVA--------------SKS 1242
Db  3019 HARAI---LIQIAKTYPQSLHFQLRTAYEDFLMLKKQQAANVLRGNSRLRENDSSSDNKS 3075

Qy  1243 --------TTTARHNAAN----------------KILKN----MCEH 1261
Db  3076 KDLSPSGSFSSVSQPNSKNGSFSSIDSSEKHQISTVKPAWELIADVTSILKTAYPLLALT 3135

Qy  1262 SNTLVQQ----AMMVSBELLIRVAILWHEWHEGLEEASRLYFGERNVKGMFEVLRPLHA 1316
Db  3136 METMVDQIHTRLKSFPEEDAYRLIV--ALLNDGLQYISRLGVVSKNT--FQ-LPMSQA 3188

Qy  1317 MMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAMDLYIHVFRRISKQLP 1376
Db  3189 NIQRFAENVLPVSVREAFLRDFVETKLDLLTYV------DKLRMW---RKKKFENILLDQRP 3239

Qy  1377 QLITSLE--LQYVSPKLLMCRDLELAVPGTY----DPNQPIIRIQSIAPSLQVITSKQR-P 1429
Db  3240 KFLHLEQCSLYLSEFQHQKFD-EVELPGQYILDKNNNDFVLRERFVPNVDLVRGHTMCY 3298

Qy  1430 RKLTLMGSNGHEFVFLLK--GHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSTQRYAV 1487
Db  3299 KRLTLRGYDGKLYPFALQYPATRHSRREERMLQLLGTFNTVLRSKIEIQNRNFSPQIPSS 3358


Qy  1488 IPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHR-----------IMLRMAPD--Y 1535
Db  3359 IPLSSHMRIIADKPSYVTMQTISDEYCKNRGMPLDYGIRFYFDRLQTGLLIQLKPRASASML 3418

Qy  1536 DHILMQKVEVFEHAVNNTAGDDLAKLLWLKSPSS----------EVWFDRRTNYT 1581
Db  3419 SNSTVEEKKQIFRQRALQLRMQLLETLNSSVFPESIYDYFYKTFERYCDFWFFRRT-FT 3477

Qy  1582 RSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREK----FPBKIPF 1637
Db  3478 TQYAYMIIMTYVFNIGGRSPQKLFTVKDSGQVMSQD----LLPSWTISNQPVFHNTEAVPF 3533

Qy  1638 RLTRMLTNAMBVTGLDGNYRITCHTWMEVLREHKDSVMAVLEAFVVDPLLNWRLMDTNTK 1697
Db  3534 RLTPPIQYLISDLGVEGLLSGLVMSIAQSLSSPTTEIKQYLSLYVRDEVFWW-------- 3585

Qy  1698 GNKRSRTRTDSYSAGQSVEILDGVELGSPAHKKTGTTVPESIHSFIGDGLVKPEALNKKA 1757
Db  3586 -------------SKQQRKPIPQGIQLPE---------TVKVNVE-----------LLFRR 3613

Qy  1758 IQIINRVRDKLTGRDFSHDDTLDVPTQVEL--LIKQATSHENLCQ 1800
Db  3614 -----------SHNVPEDLPLNQTLVDLVSQAINPQQLAQ 3646


Search completed: March  2, 2004, 20:02:11
Job time : 41 secs
```

```
Qy   1167  LTIWFDYGHPDVNEALVEGVAIQIDTWLQVIPQLIARI-----DTPRPL--VGRLIH  1218
                 ::| :||:|  : ::|   |   :|  ||:|:  :|:|: ||
Db   2251  CGLWFENDDNSKINQLLYKEIGTIPSWKFLPWVNQIASKISMEENEFQKPLQLITMKRLLY  2310

Qy   1219  QLLTDIGRYHPQALIYPLTVASKSTTTARH-NAANKILKMCEHSNTLVQQAMMVSEELI  1277
                 :|  |:|||  :| ::|| | ||  |:  |||||||
Db   2311  KLPYD-SLYSVMSLLLYBKQSNKDTNISQKIQAVKKILLEL----------------  2350

Qy   1278  RVAILWHEMWHEGLEEASRLYFGERNVKGVFEVLEPLHAWMERGPQTLKETSFNQAYGRD  1337
                 :| |                          :|| |           -AFAKKY----
Db   2351  -----------QG--------------------YDRG-------AFAKKY----  2362

Qy   1338  LMBAEQEWCRKYMKSGNVKDLITQAWDLIYTHVFRRISKQLPQLTSLELQVVSPKLLMCRDLE  1397
                 |: || | : :|:|||::|::|
Db   2363  LLPVQEFCEMSVELANLK--------FVQNTKTL-RLANLKIGQYWLKQLNMEKLP  2409

Qy   1398  L---AVPGTYDPNQPIIRIQSIAPSLQVITS-KQRPRKLTLMGSNGHEFVFLLKG-HE  1450
                 :  | || ::| | |: | | ::|| ::|| ||  ::|:||||||||||:|||  ||
Db   2410  LPTSNFTVKSSADGRKARPYIVSVNETVGITTGLSLPKIVTFNISDGTTQKALMKGSND  2469

Qy   1451  DLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALI  1510
                 |:|||  | :| :  :|:|  :: :: |   :| |: |  ::||:| ||  ||: :
Db   2470  DLRQDAIMEQVFQQVNKVLQNDKVLRNLDLGIRTYKVVPLGPKAGIIEFVANSTLHQIL  2529

Qy   1511  RDYREKKKILLNIEHRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDLAKLLWIKSPSS  1570
                 :| |:::|| :|  |||  : :|  |  |: | ::||  :||
Db   2530  SKLHTNDKITFDQARKGMKAVQTKSNEBRLKAYLK------ITNEIKPQLRNFFFDSPFDP  2584

Qy   1571  EVWFDRRTNYTRSLAVNSMVGYILGLGDRHPSNLMLDRLSGKILHIHPDGCFBVAMTREK  1630
                 |: |   : :| | :|| |:|||||||||||||||||||||||:|:||||:|: :| |::
Db   2585  LDWFEAKKTYTKGVAASSIVGYILGLGDRHLNNILLDCSTGEPIHIDLGIAFDQQKLL-P  2643

Qy   1631  FPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWR  1690
                 |||:|| | : :|: ||||| ||||||:|| || :| :|:  |  :|  |||||  ||
Db   2644  IPELVPFRILTRDIVDGFGVTGVDGLFRRSCERVYAVLRKDYVKVMCVLNILKWDPLYSWV  2703

Qy   1691  LMDTNTKGNKKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKP  1750
                 :                  |                     ||:|  -----ND
Db   2704  M----------------SPVKKYE------HLFEEHEITNF---DNVSKFISN---ND  2734

Qy   1751  EALNKKAIQIINRVRDKLITGRDFSHDETLDVPTQVELLIKQATSHENLCQCYIGWCPFW  1809
                 :   ::| :::| | ||   |:  : ||: |: ||::| | :: |: ||  ||  | |
Db   2735  RNENQESYRALKGVEEKLMG------NGLSVESSVQDLIQQATDPSNLSVIYMGWSPFY  2787
```

```
RESULT 15
ID   YAMB_SCHPO   STANDARD;   PRT;   3655 AA.
AC   Q10064;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Hypothetical protein C1F5.11c in chromosome I.
GN   SPAC1F5.11C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne R., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- SIMILARITY: STRONG, TO YEAST YHR099W.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z68136; CAA92239.1; -.
DR   PIR; T38084; T38084.
DR   GeneDB_SPombe; SPAC1F5.11c; -.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; FALSE_NEG.
DR   PROSITE; PS00916; PI3_4_KINASE_2; FALSE_NEG.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Hypothetical protein; Transferase; Kinase.
FT   DOMAIN   3324   3655   PI3K/PI4K.
SQ   SEQUENCE   3655 AA;   420774 MW;   50475E3F3FC2124A CRC64;
```

```
Query Match            4.1%;   Score 390.5;   DB 1;   Length 3655;
Best Local Similarity  18.5%;   Pred. No. 4.3e-15;
Matches  331;  Conservative 264;  Mismatches 603;  Indels 587;  Gaps 75;

Qy   357   LAAIQLFGANLDDYLHLLLPIVVLFDAPEAPLPSRKAALETVDRLTESLDFTYASRII  416
                 : :|| ||| :   |  : ::      : |::  :|   : ::
Db   2108  LPCVBCFPFKYKKDALHTLLPGFMRCF----------HKVAKEFL-----SLGSQPSGNSLN  2153

Qy   417   HPIVRTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVHRINHQRYDV--LIC  474
                 |  |  : : |   |: : ||  |  |:|:  :| :  |          ::::
Db   2154  LQIVNAVDERVSILKSMIELGCSYISYSILGDQRRWFLSALVQIEK----SSSYPICNYLL  2209

Qy   475   RIVKGYTLADEEDPLI----------YQHRMLRSGQGD---ALASGPVETGPMK  516
                 |:|:|  |  :|:|:            : :: | | |         ||:|| || |
Db   2210  EIVRGWIMNSPVPVPTVKEKAALLLKMVTFEGRFSQNEQNDLPNKYLSFIADIYEMEPYK  2269

Qy   517   KLHVSTINLQKAWGAARRVSKDWLEWLRRLSLELLK--DSSSPS---------  559
                 : :: | ::   | ::  :   :: : |  ||| |     ||:|||
Db   2270  NSEL-TFRLEAVYLLGTRVAN---------KKLKERFIKGLNSSFPSDLFSRFQFLLGSQHW  2321

Qy   560   ----------LRSCWALAQAYNPMAR-DLFNAAFVSCWSLNED------QQDELI  598
                  :: : | |  :|  |  : |: ::
Db   2322  ESLSNTYWIVQLNIFLSRCFDLNQRCQFYKKPKLFSCFSIYC-REFDELTSQAQDTEML  2380

Qy   599   RS-------IELALHSQDIAEVTQTLLNL-------AEFMEHS---------DKGPLPLR  635
                 |:       :| :| ||: ||: : ::|          |:|:
Db   2381  HNNLLKYGIIDFNQNSMLVSDPVLPVLSLQFSNSKIABYLMRDFFNASVCSFTKDFIPLC  2440
```

```
RESULT 14
ID   TEL1_YEAST      STANDARD;       PRT;   2787 AA.
AC   P38110;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Telomere length regulation protein TEL1.
GN   TEL1 OR YBL088C OR YBL0706.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=96076635; PubMed=7502586;
RA   Obermaier B., Gassenhuber J., Piravandi E., Domdey H.;
RT   "Sequence analysis of a 78.6 kb segment of the left end of
RT   Saccharomyces cerevisiae chromosome II.";
RL   Yeast 11:1103-1112(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA   Greenwell P.S., Kronmal S.L., Porter S.E., Gassenhuber J.,
RA   Obermaier B., Petes T.D.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Putative phosphatidylinositol kinase involved in
CC       controlling telomere length.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X79489; CAA56016.1; -.
DR   EMBL; Z35849; CAA84909.1; -.
DR   EMBL; U31331; AAA69802.1; -.
DR   PIR; S45416; S45416.
DR   GermOnline; 138523; -.
DR   SGD; S000184; TEL1.
DR   GO; GO:0005634; C:nucleus; IC.
DR   GO; GO:0007004; P:telomerase-dependent telomere maintenance; IMP.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Transferase; Kinase.
FT   DOMAIN    2461   2787      PI3K/PI4K.
FT   CONFLICT  1190   1190      Y -> F (IN REF. 2).
SQ   SEQUENCE   2787 AA;  321663 MW;  439B6B189E39499B CRC64;
```

```
Query Match       5.2%;  Score 494;  DB 1;  Length 2787;
Best Local Similarity  19.4%;  Pred. No. 1.2e-21;
Matches  361;  Conservative 307;  Mismatches 671;  Indels 520;  Gaps  81;

Qy    176  SLSESKSSQDS------SDYSTSEEMLVNMG------NLPLDEFYPAVSMVALMRIFRDCQS  223
                |||   :| :  ||   :  :|:    :   :::  |: :|    : |  :|::|:|
Db   1224  SLTELRSTVEKLFPTSYILSPYLFENSSVSMRYQYPLHIPL----ALGATLVQTQPAHEK 1278

Qy    224  LSHH-------TMVVQAITFIFKSLGLKCVQFLEFQVMPTFLNVIRVCDGAIREFLFQ 1166
                ||    :|  || | :  :|| |::|  : | | :  : |:| |  :  || |:
Db   1279  NNTHEFKLLFLSVITDLEKTSTYIGK--LRCARELKYLFVLYENVL--VKSSTLNFIII 1333
```

```
Qy    275  QLGMLVSFVKSHIRPYMDEIVIMREFWVNTSIQSTILLIEQIVVALGGEFKLYLPQL  334
                :|:  |: || |::|:       |:|:| :|||      |  |  : |: ||: ||:
Db   1334  RLSKFLIDTQIH-----DEVITIFSS--LLNLADKN*P----------------EIEPSLPNL 1373

Qy    335  IPHMLRVFMHDN--SPGRIVSIKLLAAIQL----------------------FGANLDDYLHLLPPI  378
                || :| :|:|:  ||:| |: ::        :|                  |: :|:|:|:|||:
Db   1374  FCKIF-IYLRENKQLSPSFQQAIKLLEHRDLIKKTWKYFLDAIFGNIVQDDIY-------EN 1428

Qy    379  VKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIVRTLDQSPELRSTAMDTLS  438
                |::| :  | |:|:|| |   : :|||  |  :   :|   |:|: ::|    :   |:|:
Db   1429  TELLDASDCGVD---------DVVLVSLLFS-YARR--PVASKIGCSLS-KAAAINILK 1474

Qy    439  SLVFQ--LGKKYQIFIPMVNKVLVRHRINHQR  468
                |:: |  |:|:|    |: : :| : :| | :|:
Db   1475  HHVPKEYLSKNFKLWFAALSRRILQQEVQRERSTNFNEVELKNPEMVVRHPEQPHMIYQ 1534

Qy    469  ----------YD------------VLICRIVKGYTLADEEEDPLIYQHRMLRSGGQDALASGP  509
                              :|   |    : | | : :||:|||::::::| | : |: |
Db   1535  RISTFNKEAELYDSTEVFPISBCILTYIVGYSIGNSESE-PCFRDNIMNENKD---------------- 1586

Qy    510  VETGPMKKLHVSTI-NLQKAWGAARRVSKDDWLE------------------WLRRLSLELLKDSS--SPS  559
                |: |:||:|  |:|  ||: |   | | :  |::|                   |  |||:|| :|   :||
Db   1587  -KVAPLDKDVLNAIYPLANNFGWESFIC-DTYLSVNEPYNCWLSKFARSLIHQISFNIPP 1644

Qy    560  LRSCWALAQAYNPMAR-------DLFNAAFV----DLFNAAFV----SC--WSELNEDQQDELIRSIEHLALTSQD  609
                |||  ||| :|:|:|:         | | |:     | | |:     ||  |: |:||:||||||:|:||::|:|
Db   1645  IVCLYPLCKGSTAFCELVLTDLFFLSTTYDPKSCLNWS-----------NRIFTQIAMLLHVKD 1697

Qy    610  IAEVTQTLLNLAEFMEHSDKGPIPLRDDNGIVLLGERAAK----------------------CRAYA  654
                ||:   | :|  :||:||:||:| |::||:|:| | |:|:                      || | :
Db   1698  SEIKLKMLFNVIK---------------------MIRMGSRCKERNCLRIYSSLDLQEICQISL 1740

Qy    655  KALHYK--ELEFQKGPTPAILESLISINNKLQQ----------------------PRAAAGVLBYAMKH  699
                |||: :  || | |  | | |: : :|:|::| :                       |  |  |: ||
Db   1741  KIKEFKFGYLLFEEMNMPNIREMNINTLQKIYECINDGDFLAGLPVPHSIEGVLNSINRI 1800

Qy    700  FGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCE  759
                |  |||: |||:||:|:|:|| |  | |  |:||:|: ||: | | || || |: :  ||
Db   1801  DSD----------TWKRFLFNNADFDANYTTSLEEEKB--------SLIKATEBDSGFYG--------- 1841

Qy    760  KWTIVNDETQAKMARMAAAAAWG--LGQWDSMEBYTCMIPR--DTHDGAFYYRAV-ALHQD  815
                || |:|:| |:|  :  :   |   ||:|||:| |  |:|:  : : |  | | : | |||
Db   1842  ---LTSLLESRLSGSSDVYKWNLELGDW------------KLLTPKVVDSKAKGLYYAIKNLPQD 1891

Qy    816  LFSLAQQCIDKARDLLDAELTAMAGESYRAYGAMVSCHMLSELE------------BVIQYK  866
                ||| ||  ||:| |:: | :  || :| | ||  : ||:| :|              |: ||
Db   1892  V-GFAEKSLEKSL----------LTIFDSRQ------------HFISQTEWMDTLNALIEPIKIA 1933

Qy    867  LVPERREIIRQIWERLQGCQRIVEDWQKILMVRSLVVSPHE-----------DMRTWLKYA-  917
                |  |::  ||:|| || | :|:|   ||:|  || : |:|            ::| :||
Db   1934  AIPQD---------VTSFPQTLM--SIMKADKERLNTIDFYDHKTTLKGRH 1973

Qy    918  ----------SI------CGKSGRL-------------------------ALAHKTLVLLIGVDPS  942
                          :|      | |:| |                         | | |  ||:|:|:
Db   1974  TLMNVLSRNSLDRNVKCSKYLRLGSIIQLANTVQLAIANGAPQDALRNATILMSKTVKNIA 2033

Qy    943  RQLDHP--LPTVHPQVTYAYMKNMWKSARKIDAFQEMQHFVQTMQQQAQHAIATEDQQHK  1000
                | | |   |:|  | | | : :   |: | ||  ::|: | | |:: | |: |:| |::|
Db   2034  KLYDDFSVVSQIEKLASFTSANALWESREYKAPVMIRDLLAQNEKNISESILYDD--FK 2091

Qy   1001  QELHKLMARCFLKLGEWQLNLQGINESTIPKVLQYYSAATEHDRSWYKAWEJAWAVMNFE-  1059
                ::|| ::  | |::  : ::|         :|  : :  :|        :|   :| | |
Db   2092  LLINVPMDQIKARLVKWS--------SESRLEPAAAIYE---------KIIVNWDI-NVED 2134

Qy   1060  -----AVLHYKHQNCARDEKKLRHASGANITNATTAATTASTEGSNS---ESE  1112
                     | ||:|:|  || :  |:        : | |::  | |: ||   |||
Db   2135  HBSCSDVFTLGSFLDEQAQKLR--SNGEIEDREHRSYIGKSTLKALELIYKNTKLPENE 2192

Qy   1113  AESTENSPTPSPLQKKVTEDLSKLLM----YTVPAVQGFFRSISLSR--GNNLQDTLRV  1166
                | | :||| || ::  :|| |:|:|:      | |  || |: :|:|   |:|| |:|
Db   2193  RKDAKRHYNRVLLQYNRDSEVLKALLLQKEKFIWHALEFYLNTIVFSNRYDNDIID--KF 2250
```

```
Db   2143 LAKLVINTEEVFRPYAKHWLSPLLQLAVCE------NNREGLHYMMVEIVATILSWTGLA 2196

Qy     73 MRKWV--DELFI-IIMDMLQDSLLAKRQVALWTLGQLVASTGYVVE-------PYRKY    121
              ::    :::                  :::
Db   2197 TPTGVPKDEVLANRLLRFLMKHVFHPKRAVFRHNL-EIIKI---LVECWKECLSIFYRL- 2251

Qy    122 PTILEVLLNFLKTEQNQGTRREAIRVLGLLGA--LDPYKHKVNIGMIIQSRDASAVSLSE  179
                                           :::
Db   2252 ------IFEKFSHKDDPNSKDNSVGIQLLGIVIANNLPPY--------DPNCDITSAMYFE 2297

Qy    180 SKSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMV-------ALMRIFRDQ            222
             :                         :::
Db   2298 A---------LVNNMSFVXYKEVYAAAAEVLGLLLQYITERKHVIAELVCELVIKQ      2344

Qy    223 SLSHHHTM-------------VVQAITP------IFKSLGLKCV----               247
                   :::                            :::
Db   2345 LKQHQNTMEDKFIVCLNKIAKGFPPLADRFLNALFFLLPKPHGVMKTLCLEVVLCRAEEI  2404

Qy    248 --QFLPQVMPTFLNVIR--------VCDGAIREF------LFQQLGMLVSFVKSHIRP     289
                                                   :::
Db   2405 TGLYLQLKSKDFLQVMRHRDDERQKVCLDIVYKMVAKLKPIELRRLLNPVVEFV-SHPSP 2463

Qy    290 ---------------YMDE-------------------IVTLMREFWVM            304

Db   2464 TCREQMYNILMWIHDNYRDQSSQNDEDSQEIFKLAKDVLIQGLIDENVGLQLIIRNFWSH 2523

Qy    305 NTSIQSTII--LL--------IEQIVALGGEFKLYIPLQLIPEMLR-VFMHDNSPGRIV   352
                         ::    :::
Db   2524 ETRLPSNTLDRLLALNSLYSPKIEVHFLSLATNFLLEMTRMSPDYLNPIFEH-------- 2575

Qy    353 SIKLLAAIQLFGANLDDYL-------HLLLPPIVKLFDAP---------EAPLPSRKAA   395
                                                :::
Db   2576 -------PLSECEFQEYTIDDPDWRFRSTVLFTPMFIETQASPSILHTCTQEGPLSDQRQK 2627

Qy    396 LETVDRLITESLDFTDYASRIIHPIVRTLDQSPELRSTAMDTLSSLVFQLGKKYQIFI-P 453
                    :::                                 :::
Db   2628 PGQVRATQQQYDFT--------PTQASVERS-------SFDWLTGSSIDLLADHTVFSSET 2673

Qy    454 MVNKVLVRHRINHQRYDVLICRIV--------KGYTLADEEEDPLIYQHRMLRSG------QGDA 504
                         :::
Db   2674 LSSSLLFSHK-RITEKSQRMSCKSVGPDFGCTKKLGLPDDEVD----------NQVKSGTPSQADI 2726

Qy    505 LASGPVETGPMKKLLHVSTINLQKAWGAARRVSKDDWLEW---------LRRLSLEL    551
             ||        :::                            :::
Db   2727 LR--LRRRFLKDREKLLSLLYAKRGLMEQKLEKDIKSEFKMKQDAQVVLYRSYRHGDLPD 2783

Qy    552 LKDSSSPSLRSCWALLAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRSIELALTSQDIA 611
              :::                                   :::
Db   2784 IQIQHSGLITPLQAVAQKDPIIIAKQLFSSLFSGILKEMNKFK----TTSEKNIITQNLL 2838

Qy    612 B------VTQTLLNLAFM--------EHSDKGPL--------PLRDDNGIVLLGRRA    647
                   :::                         :::
Db   2839 QDFNRFLNTTFLFPPPFVSCIQEISCQHPDFLTLDPASVRVGCLASLQQPGGIRLLEE---- 2896

Qy    648 AKCRAYAKALHYKELEFQKGPTPAILE--SLLISNNKLQQPEAAAGVLEYAMKHFGELEI  705
                    :::                                :::
Db   2897 ALLRLLMPKEPPTKRVRGKTCLPPDVLRWMELAKLYRSIGEYDVLRGLFSSELGTTQDTQ- 2955

Qy    706 QATWYEKLHEWEDALVAYDKKMD--------TNKDPDELMLGRMRCLEALGEWGQLHQQ   756
                    :::                              :::
Db   2956 NALLAEARSDYCCQAAKLYDEALNKLEWDGEPTEARKEFWELASLDCYNNLSKNKEL--E  3013

Qy    757 CCEKWTLVNDETQAKMARMAAAAAWGLGQWDS--MEEYTCMLPR----------      798
                    :::                            :::
Db   3014 YCSTVNIVS-ENSLDLSKM--------WSBPFYQETYLPVVIRSKLKLLQGEGNQSL    3062

Qy    799 ----DTHDGAFYRAVLALHQ----DLFSLAQQCIDKAR------DLLDAELTAMAGES   842
                    :::                                 :::
Db   3063 LTFVDEAMNKELQKTVLELQYSQELSLLYILQDDIDRAFYIKNGIQIFMQNYSSIDVLL  3122

Qy    843 YSRAVGAMVSCHMLSELEEVIQY---------KLVPERREIIRQIWWERLGQCQRIVED--- 892
              :::                                    :::
Db   3123 YRSRLAKLQSVQTLAEIEEFLSFICKHGDLSSLGPLRR--LLKTWTSRYPD---VVTDPM  3177
```

```
Qy    893 --WQKILMVRSLVVS-------------------------PHEDMRTWLKYASLC       920
              :::
Db   3178 HIWDDIITNRCFFLSKIEBRLTAPSGDHSMSVDEDEESIDREVYEPKEDVRCMLQSCRFT 3237

Qy    921 GKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKS--ARKIDAFQHMQ   978
              :::                                         :::
Db   3238 MKMKMIRSAWKQSNFSLSMKLLKEMHKESKT--------REIWRVQWLHSYSQLNHCR   3287

Qy    979 HFVQTMQQQAQCHAIAT-----EDQQHKQELHK-LMARCFLK-------------LGEWQLN 1020
              :::                                           :::
Db   3288 SHTQSPREQVLNTLKTITLLDBSDISNYLNKNIQASCDQSLLGGTCRIMADALSREPAC  3347

Qy   1021 LQGINESTIPKVLQYYSAATEHDRS------WYKAWHAWAVMNFEAVLHYKHQNQARDEKK 1075
              :::                                              :::
Db   3348 LSDLEERNKVNSILTLSGSNAENTEVVITGLYQRAFH--------ELSKAVQSAEEET   3396

Qy   1076 KLR--HASGANITNATTAATTAATTASTEGSNSESEAESTENSPTPSPLQKKVTED     1132
              :::                                              :::
Db   3397 QLSCWGHEAAAAERAHAYMTLVGFCDQLRKVEESASQKTSAE-MEAYPA----LVVEK   3449

Qy   1133 LSKTLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQI  1192
              :::                                              :::
Db   3450 MLRALKLNSSEARIKFPRLLLQIIEQYS-EETLNIMT--------KEISSIPCWQF    3495

Qy   1193 DTWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKS----TTTARH 1248
              :::                                              :::
Db   3496 IGW---ISHMALLDKEEAIA---VQHTVBEIADNVPQALIYPIISSESYSFKNTSSGH   3549

Qy   1249 NAANKI----LKDMCEHSNTLVQQAMMVSEBLIRVAILWHEMWHEGLEEASRLYFGERNV 1304
              :::                                              :::
Db   3550 N--NKAFVERIKSKLDHGE-VIHSFINALDQLSNPDLLFKDWVSDTKDELGKNPVNKKNI  3606

Qy   1305 KGMFE----VLEPLHAMMERGPQTLKETSFNQAYGRDL----------MEAQEWCR-    1346
              :::                                         :::
Db   3607 EKLYERMYAALGDIRA-PGLGPFFRR--FIQAFGKEFVKSFGNGGSKLLTMKVDDFCKI  3662

Qy   1347 ------KYMKS---GNVKDLTQAWDLIYHVFRRISKQLPQLTSLELQYVSPKLLMCRDL  1396
              :::
Db   3663 TGSLLVRMKKDSKLPGNLKEYS-------------------------PWMSBFKAQFLKN-- 3697

Qy   1397 ELAVPGTYD-PNQPI----IRIQSIAPSLQVTSKQRPRKLTMGSNGHEFVFLLKGHED  1451
              ||            :::                              :::
Db   3698 ELEIPSQYDGKSKPLPEYHVRISGFDERVKVMLSLRKPKRIVIRGHDRKEYPFLVKGGED 3757

Qy   1452 LRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTL-HALI 1510
              :::                                              :::
Db   3758 LRQDQRIBQIFEVMNAILSQDAACSQRNMQLRTYRVVPMTSRIGLIIEWIENTMTLKDLL  3817

Qy   1511 RDYREKKKLLLNIEHRIMLRMAPD--------YDHLTLMQKVEVFEHAVNN          1553
              :::                                   :::
Db   3818 SNMSQBEKVANNSDPKAPIRDYKDWLMKVSGKSDAGAVLMYSRANRTETVVAFRRRESQ  3877

Qy   1554 TAGDDLIAKLLWLKSPSSEVWFDRRTNYTRSLAVNSMVGYILGLGDRHPSNLMLDRLSGKI 1613
              :::                                              :::
Db   3878 VPPDLLKRAFVKMSTSPEAFLALRSHPASSHALLCISEWLLGIGDRHLNNFMVAMETGSV  3937

Qy   1614 LHIDFGDCFEVAMTREKFPEKIPFRLTRMLTN---AMEVTGLDGNYRITCHTWMEVLREH 1670
              :::                                              :::
Db   3938 IGIDFGHAPGSATQFLPVELMPPRLTRQFVSLMLPVKKETGL-------MCTVMVHALRAF 3991

Qy   1671 KDS---VNAVLEAFVYDPLLNWRLMDTNTKGNKSRTRTDSYSAGQSVEILDGVELG-EP  1726
              :::                                              :::
Db   3992 RSCAGLLTDTMEIFVKEP2SPDW----------KSFEQTMLRKGGSWIQEINVTEKNWYP  4040

Qy   1727 AEK------KTGTTVPESI--HSFIGDGLVKPEALNK-KAIQIINRVRDKLTGRDFSHD  1776
              :::                                              :::
Db   4041 QHKIRYAKRKLAGANPAVITCDELYLGH--------EASSAFRSYTAVARGNRDYNIRAQPE 4095

Qy   1777 DTLDVPTQVELLIKQATSHENLCQCYIGWCPF                           1808

Db   4096 SGLSBETQVKCLVDQATDPNILGRTWEGWEFW                           4127
```

```
Qy     1790 KQATSHENLCQCYIGWCPF 1808
            ||||  |||||.|  || ||.|
Db     4109 DQATDPNILGRTWEGWHEPW 4127
```

RESULT 13
PRKD_MOUSE
ID   PRKD_MOUSE   STANDARD;   PRT;   4128 AA.
AC   P97313; 081187; P97928; Q9Z341;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   DNA-dependent protein kinase catalytic subunit (EC 2.7.1.37) (DNA-
DE   PKcs) (P460).
GN   PRKDC OR XRCC7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANT SCID.
RC   STRAIN=C.B17; TISSUE=Fibroblast, and Leukocyte;
RX   MEDLINE=97225971; PubMed=9122213;
RA   Araki R., Fujimori A., Hamatani K., Mita K., Saito T., Mori M.,
RA   Fukumura R., Morimyo M., Muto M., Itoh M., Tatsumi K., Abe M.;
RT   "Nonsense mutation at Tyr-4046 in the DNA-dependent protein kinase
RT   catalytic subunit of severe combined immune deficiency mice.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:2438-2443(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98250757; PubMed=9582343;
RA   Fukumura R., Araki R., Fujimori A., Mori M., Saito T., Watanabe F.,
RA   Sarashi M., Itsukaichi H., Eguch-Kasai K., Sato K., Tatsumi K.,
RA   Abe M.;
RT   "Murine cell line SX9 bearing a mutation in the DNA-PKcs gene exhibits
RT   aberrant V(D)J recombination not only in the coding joint but also in
RT   the signal joint.";
RL   J. Biol. Chem. 273:13058-13064(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=97480731; PubMed=9339376;
RA   Fujimori A., Araki R., Fukumura R., Saito T., Mori M., Mita K.,
RA   Tatsumi K., Abe M.;
RT   "The murine DNA-PKcs gene consists of 86 exons dispersed in more than
RT   250 kb.";
RL   Genomics 45:194-199(1997).
RN   [4]
RP   SEQUENCE OF 1-26 FROM N.A.
RX   MEDLINE=98384545; PubMed=9716665;
RA   Saito T., Matsuda Y., Ishii H., Watanabe F., Mori M., Hayashi A.,
RA   Araki R., Fujimori A., Fukumura R., Morimyo M., Tatsumi K., Hori T.,
RA   Abe M.;
RT   "Mouse Cdc21 only 0.5 kb upstream from DNA-PKcs in a head-to-head
RT   organization: an implication of co-evolution of ATM family members and
RT   cell cycle regulating genes.";
RL   Mamm. Genome 9:769-772(1998).
RN   [5]
RP   SEQUENCE OF 3615-4128 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Leukocyte;
RX   MEDLINE=97045026; PubMed=8881030;
RA   Hamatani K., Matsuda Y., Araki R., Itoh M., Abe M.;
RT   "Cloning and chromosomal mapping of the mouse DNA-dependent protein
RT   kinase gene.";
RL   Immunogenetics 45:1-5(1996).
RN   [6]
RP   SEQUENCE OF 3680-4128 FROM N.A.
RX   MEDLINE=96413638; PubMed=8816792;
RA   Blunt T., Gell D., Fox M., Taccioli G.E., Lehmann A.R., Jackson S.P.,
RA   Jeggo P.A.;
RT   "Identification of a nonsense mutation in the carboxyl-terminal region
RT   of DNA-dependent protein kinase catalytic subunit in the scid mouse.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:10285-10290(1996).

RN   [7]
RP   SEQUENCE OF 3839-4128 FROM N.A.
RC   STRAIN=C.B17;
RX   MEDLINE=96413301; PubMed=8816463;
RA   Danska J.S., Holland D.P., Mariathasan S., Williams K.M., Guidos C.J.;
RT   "Biochemical and genetic defects in the DNA-dependent protein kinase
RT   in murine scid lymphocytes.";
RL   Mol. Cell. Biol. 16:5507-5517(1996).
-!- FUNCTION: SER/THR KINASE INVOLVED IN DNA DOUBLE-STRANDED BREAK
CC   REPAIR, V(D)J RECOMBINATION AND MODULATION OF TRANSCRIPTION. MUST
CC   BE BOUND TO DNA TO EXPRESS ITS CATALYTIC PROPERTIES.
-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
-!- SUBUNIT: FORMS A COMPLEX WITH THE HETERODIMER KU70/KU80
CC   AUTOANTIGEN. INTERACTS WITH DNA-PKCS INTERACTING PROTEIN (KIP)
CC   WITH THE REGION UPSTREAM THE KINASE DOMAIN.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- DISEASE: Defects in PRKDC are the cause of severe combined immune
CC   deficiency (SCID) which is characterized by a lack of mature
CC   functional lymphocytes and a high susceptibility to lethal
CC   opportunistic infections if not chronically treated with
CC   antibiotics. The lack of B- and T-cell immunity ressembles severe
CC   combined immunodeficiency syndrome in human infants.
-!- SIMILARITY: Belongs to the PI3/PI4-kinase family.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; D87521; BAA19566.1; -.
DR   EMBL; AB007544; BAA28873.1; -.
DR   EMBL; AB030734; BAB91149.1; -.
DR   EMBL; AB011543; BAA28875.1; -.
DR   EMBL; AB000629; BAA34640.1; -.
DR   EMBL; D83786; BAA12115.1; -.
DR   EMBL; U78157; AAB36939.1; -.
DR   EMBL; U78158; AAB36940.1; -.
DR   PIR; JC6306; JC6306.
DR   MGD; MGI:104779; Prkdc.
DR   InterPro; IPR009938; FAT.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; FALSE_NEG.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Transferase; Serine/threonine-protein kinase; DNA repair;
KW   Nuclear protein; Disease mutation; SCID; Polymorphism;
KW   Phosphorylation.
FT   DOMAIN   3748   4128   PI3K/PI4K.
FT   MOD_RES   2605   2605   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   MOD_RES   2608   2608   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   MOD_RES   2634   2634   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   MOD_RES   2643   2643   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   VARIANT   2140   2140   C -> R.
FT   VARIANT   3191   3191   L -> P.
FT   VARIANT   4046   4128   MISSING (IN SCID).
FT   CONFLICT   3844   3844   M -> V (IN REF. 6).
SQ   SEQUENCE   4128 AA; 471383 MW; C2D7368DB50AEE0 CRC64;

Query Match   5.9%; Score 552.5; DB 1; Length 4128;
Best Local Similarity   18.8%; Pred. No. 5.5e-25;
Matches 411; Conservative 356; Mismatches 818; Indels 607; Gaps 94;

Qy   17 LGHLVSNAPRLIRPY----MEPILKALILKLKDPDPDNPGVINNVLATIGELAQVSGLE 72
             ||| ||| || .|: : :: .|| :|: . :   :: .||
```

```
Matches  433;  Conservative  364;  Mismatches  798;  Indels  584;  Gaps  101;

Qy    17  LGHLVSNAPRLLIRPYMEPIIKALILKLKDPDPDNPGVINNVLATIGELAQVSGLEMRKW      76
          ||| : ||   |  :|    :: :   :::|| :    |  : :  :| :   |
Db  2146  LAKLVINTEEVFRPYAKHWLSPL-LQLAASENNGBGIHYMVVEIVATILSWTGLATPTG    2204

Qy    77  V--DELFI-IIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVE------PYTRKYPTLL     125
          |  |:|:  ||| |       :| :|||  :       ||| ||
Db  2205  VPKDEVLANRLLNFLMKHVHPKRAVFRHNL-EIIKT--LVECWKDCLSIPYRL-----     2255

Qy   126  EVLLNFLKTEQNGTRREAIRVLGLLGA--LDPYKHKVNIGMLDQSRDASAVSLSESKSS     183
          : : |    ::: :   |:|:  |  :
Db  2256  -IFPKFSGKDPNSKDNSVGIQLLGIVMANDLPPYDPQCGI----------            2294

Qy   184  QDSSDYSTSEMLVNMGNLFLDEFYPAVSMV----------ALMRIFRDQSLSH          226
           ::  :| |    ||   || |                     |||:|||:||:
Db  2295  -QSSEYFQA-LVNNMSFVRYKEVYAAAAEVLGLILRYMERKNILEESLCELVAKQLKQH    2352

Qy   227  HHTM----VVQATTFIFKSLGLKCVQFLPQVMPTFLNVIR----VC-DGAIREF         271
          ::|     ||| :  :  | :  | |||   : |:|||
Db  2353  QNTMEDKFIVCLNKVTKSPPPLADRFMNAVFFLLPKFHGVLKTLCLEVVLCRVEG-MTEL   2411

Qy   272  LFQOLGMLVSFVKSHIRPYMDEIVTLMREFWVMNTSIQSTIILLEQIVVALGGEFKLYL    331
          :|| ||                |   :||  | |:    :||||||:|:||||||||:|
Db  2412  YFQ-------LKS------KDFVQVMRH---RDDERQKVCLDII---------YKM-M    2443

Qy   332  PQLIPEMLR----------VFMHDN-------SPGRIVSIKLLAA                  359
          | |: |:||         |:| ||       |||:|:||||:|
Db  2444  PKIkKPVELRELNLPVVEFVSHPSTTCREQMYNIIMWIHDNYRDPESETDNDSQEIFKLAK  2503

Qy   360  IQLFGANLDDYIHLLLPPIVKLFDAPEAPLPSRKAALBTVDRLITE-----------     404
          | ||| :|: |||||||||:||| || | |||||||| ||||:
Db  2504  DVLIQGLIDENPGLQL--IIRNFWSHETRLPS------NTLDRLLALNSLYSPKIEVHFLS  2556

Qy   405  -----SLDFT-----DYASRII-HPIVR-------TLDQSPELRSTAM-------DTL    437
               |||:      ||||:|| |||:|       ||||||||||:||       |||
Db  2557  LATNFLLEMTSMSPDYPNPMFEHPLSBCEFQEYTIDSDWRFRSTVLTMFVETQASQGTL    2616

Qy   438  SSLVFQ--LGKKYQIFIPMVNKLVRHRINHQRYDVLICRIVKGYTLAD-----------  484
          ::|:|   |||||| | |||||||:||| || ||||| |||||| |||
Db  2617  QTRTQEGSLSARW----PVAGQI---RATQQQHDFTLTQTADGRSSFDWLTGSSTDPLV   2668

Qy   485  ---EEEDPLIYQHRMLRSGQGDALASGPVET-GP--MKKLHVSTINL-QKAWGAARRV    535
             |||||||||||:|| |||||||:|| |:  ||  ||||||||||| |||||||||||
Db  2669  DHTSPSSDSLLFAHK------RSERLQRAPLKSVGPDFGKKRLGLPGDEVDNKVKGAAGRT  2723

Qy   536  SKDDWLEWLRRL----------SLELLKDSSSPSLRS------                  562
          |:|||||||||           ||||||||||||||
Db  2724  ---DLLRLRRRFMRDQEKLSLMYARKGVAEQKREKEIKSELKMKQDAQVVLYRSYRHGDL  2780

Qy   563  ---------CWALAQAYNPMARDLFNAAFVSCWKSELNE------DQQDELIRSI      601
                   |||||||||| |||||||| | ||| |||        |||||||||:|
Db  2781  PDIQIKHSSLIITPLQOAVAQRDPIIAKQLFSSLLFSGILKEMDKFKTLSEKNNITQKLLQDF  2840

Qy   602  ELALTS--------QDIAEVTQTLINLLAEFMEHSDKGPL--PLRDDNGIVLL         643
          | | :|        ||| || ||| | | | ::|:  ||  |||||||||:||
Db  2841  NRFLNTTFSFPPPFVSCIQDISCQHAALLSL--------DPAAVSAGCLASLQQPVGIRLL  2893

Qy   644  GERAAKCRAYAKALHYKELEFQKGPTPAILE--SLISINNKLQQPEAAAGVL-EYAMKH   699
          ||||||||||||||||||  |||||| | |   |||||||||||| || ||| |||||:
Db  2894  EE---ALLRLLPAELPAKPVRGKARLPPDVLRWVELAKLYRSIGEYDVLRGIFTSEIGTKQ  2951

Qy   700  FGB----LEIQATWYEKLHEWEDALVAYD---KKMDTNKDDPELMLGRMRCLEALGEWG   751
          |||    ||| | ||:|||||||||| |    |||||||||||||||||:|||||||||
Db  2952  ITQSALLAEARSDYSEAAKQYDEALNKQDWVDGEPTEAEKDFWE--LASLDCYNHLAEWK  3009

Qy   752  QLHQQCCEKWTLVNDE---------TQAKMARMAAAAAWGLGQWDSMEEYTCMIPR      798
          |||||||||| | |           |||:||||||||||||||||||||||| |||
Db  3010  SL--EYCSTASIDSENPPDLNKIWSEPFYQETYLPYMIRSKLKLLQGEADQSLLTFIDK   3067

Qy   799  DTHDGAFYRAVLALHQ-----DLFSLAQQCIDKARDLLDAELTAMAGESYRAYGA-----  849
          :  :|  :||:|||| |     ||||||||||||||||||||||||||||||||
Db  3068  AMH-GELQKAILELHYSQELSLLYILQDDVDRAKYYIQNGI----QSFMQNYSSIDVLL   3121

Qy   850  ------MVSCHMLSLSEVIQY-----                    KIVPERREIIRQIWERLQGCR-IVEDW   893
                :| |:|||||:|:|                          |:|| |:|:|||||||||| :|||
Db  3122  HQSRLTKLQSVQALTFIQEFISFISKQGNLSSQVPLKR--LLNTWTNRYPDAKMDPMNIW   3179

Qy   894  QKILMVRSLVVSPHEDMRIWLKYASLCGKSGRLALAHKTLVLLVDLDPSRQLDH---PLP   950
          ||: |  |:   ||:|:| |||||||||||||||||||||||||||||||||||   |||
Db  3180  DDITTNRCFFLSKIIEKLT------------------PLPEDNSMNVDQDGDPSD   3216

Qy   951  TVIHPQVTYAYMKMWKSARKIDAFQEMQHFVQIMQQAQHAIATEDQQHKQELHK------  1005
          |::| ||:|||||||||||||:||::||| |||||||||||||||||||||||
Db  3217  RMEVQEQEDISSLIRSCK-------FSMKMKMIDSARKQNNFSLA---MKLLKELHKESKTR  3269

Qy  1006  -----LMARCFIKLGEWQLNLQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEA  1060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3270  DDWLVSWVQSYCRLSHCRSRSQGCSEQVL-TVLKTVSLLDENNVSSYLS--------KN   3319

Qy  1061  VLHYKHQN---------QARDEKKGLRHASGANITNATTAAT--T                1094
          :|:|||   |:||::|||:||| ||||||||||||||||||:  |
Db  3320  ILAFRDQNILLGTTYRIIANALSSEPACLAEIEBDKARRIELSLSGSSSEDSEKVIAGLYQ  3379

Qy  1095  AATATTTASTEGSNSESEAESTENSPTPSPLQKKVT------EDLSK------TL       1137
          |||||||||||||||||||||||||||||||||||      ||||       |:
Db  3380  RAFQHLSEAVQAAEBEAQPPSWSCGPAAGVIDAYMTLADFCDQQLRKEENASVIDSAEL   3439

Qy  1138  LMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWPDYGHWPDVNEAL----VEGVKAIQID  1193
          |||||||:|||||||||||||||||||||||||||||||||||||    |:|||||:||
Db  3440  QATPALVVEKMLKALKL---NSNEARLKFPRLLQIIRYPRETLSLMTKEISSVPCWQFI   3496

Qy  1194  TWLQVIPQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVAS-----KSTTTARH  1248
          :|| |||||||||||||||||||||||||||||||||||||||||:      |||||||
Db  3497  SW---ISHMVALLDKDQAVA---VQHSVEITDNYPQAIVIPFIISSESYSFKDTSTGHK   3550

Qy  1249  NAANKI-LKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGM  1307
          |:|:|  ||:||||||||||||||||||||||||||||||||||||||||||||:|||||
Db  3551  NKEFVARIKSKLDQGG-VIQDFINALDQLSNPELLFKDWSNDVRAELAKTPVNKKNIEKM  3609

Qy  1308  FEVL------EP---LHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLT  1358
          :|:|      ||   |||||||||||:|||||||||||||||||||||||||||||||||
Db  3610  YERMYAALGDPKAPGLGAFRRKFIQTFGK-EPDKHFGKG---GSKLLR--MKLSDFNDIT  3663

Qy  1359  QAWDLYHVFRR----ISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYD-PNQPI----  1410
          |||||||||||    ||||||||::||:|||||||||||||||||| ||  ||||
Db  3664  NMLLLKMNKDSKPPGNLKECSPWMSDFKVEFLRN----------ELEIPGQYDGRGKPLPEY  3715

Qy  1411  -IRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVELLKGHEDLRQDERVMQLFGLVNTLL  1469
           |||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db  3716  HVRIAGFDERVTVMASLRRPKRIIIRGHDEREHPFLVKGGRDLRQDQRVEQLFQVMNGII  3775

Qy  1470  ANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRD--YRKKKILLN------  1522
          |:||||||||||||||||||||||||||||||||||||||||   ||||||||
Db  3776  AQDSACSQRALQLRTYSVVPMTSRLGLIEWLENTVTLKDLLLNTMSQEEKAAYLSDPRAP  3835

Qy  1523  -IBHRIML-RMAPDYD--HLTLMQK-----VEVEHAVNNTAGDDLAKLLWLKSPSSE     1571
           |:|||:| ||||||   |||:|||     |||||||||||||||||||||||||||
Db  3836  PCEYKDWLTKMSGKHDVGAYYMLMYKGANRTEVTSFRKGRSKVPADLLKRAFVRMSTSPE  3895

Qy  1572  VWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKF  1631
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3896  AFFALRSHFASSHALICISHWILGGDRHLNNFMVAMETGGVIGDFGHAFGSATQFLPV   3955

Qy  1632  PEKIPFRLIRMLIN--AMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLN   1688
          ||||||||||||:   ||||||||||||||||||||||||||||||||||||||||||
Db  3956  PELMPFRLTRQFINLMLPMKETGL--MYSIMVH-ALRAFRSPDGLLINTMDVFVKEPSPD  4012

Qy  1689  WRLMDTN--TKG-----NKRSRTRTDSYSAGQSVEILDGVELGEPAHK              1729
          |||||:   |||     |||||||||||||||||||||||||||||||
Db  4013  WKNFEQKMLKKGGSWIQEINVABKNWYPRQKICYAKRKLAGANPAVITCDELLLG---HE  4069

Qy  1730  KIVGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLLVPTQVELLI  1789
          |  |:|||| |:   | |:|:|   |||||||||||||||||||||||||:: |||:
Db  4070  KA-----PAFRDYVA------VARGSKDHNI---RAQBPSSG--------LSEETQVKCLM  4108
```

RL   Cell 82:849-856(1995).
RN   [2]
RP   REVISIONS, AND ALTERNATIVE SPLICING.
RA   Gell D., Anderson C.W.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
RA   Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA   Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 1-1689 FROM N.A.
RX   MEDLINE=21312114; PubMed=11418067;
RA   Anderson C.W., Dunn J.J., Freimuth P.I., Galloway A.M.,
RA   Allalunis-Turner M.J.;
RT   "Frameshift mutation in PRKDC, the gene for DNA-PKcs, in the DNA
RT   repair-defective, human, glioma-derived cell line M059J.";
RL   Radiat. Res. 156:2-9(2001).
RN   [5]
RP   SEQUENCE OF 1-49 FROM N.A.
RX   MEDLINE=97430835; PubMed=9284934;
RA   Ladenburger E.M., Fackelmayer F.O., Hameister H., Knippers R.;
RT   "MCM4 and PRKDC, human genes encoding proteins MCM4 and DNA-PKcs, are
RT   close neighbours located on chromosome 8q12-->q13.";
RL   Cytogenet. Cell Genet. 77:268-270(1997).
RN   [6]
RP   SEQUENCE OF 1789-2203 FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95365397; PubMed=7638222;
RA   Sipley J.D., Menninger J.C., Hartley K.O., Ward D.C., Jackson S.P.,
RA   Anderson C.W.;
RT   "Gene for the catalytic subunit of the human DNA-activated protein
RT   kinase maps to the site of the XRCC7 gene on chromosome 8.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:7515-7519(1995).
RN   [7]
RP   SEQUENCE OF 2255-2335 FROM N.A.
RC   TISSUE=Placenta;
RA   Abe M.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 3199-4128 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal lung;
RX   MEDLINE=96062265; PubMed=7594449;
RA   Poltoratsky V.P., Shi X., York J.D., Lieber M.R., Carter T.H.;
RT   "Human DNA-activated protein kinase (DNA-PK) is homologous to
RT   phosphatidylinositol kinases.";
RL   J. Immunol. 155:4529-4533(1995).
RN   [9]
RP   SEQUENCE OF 3250-4128 FROM N.A. (ISOFORM 1).
RA   Anderson C.W., Dunn J.J., Freimuth P.I.;
RT   "Sequence of the 3' segment (exons 70-86) of PRKDC, the gene for human
RT   DNA-PKcs.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=97074683; PubMed=8917110;
RA   Connelly M.A., Zhang H., Kieleczawa J., Anderson C.W.;
RT   "Alternate splice-site utilization in the gene for the catalytic
RT   subunit of the DNA-activated protein kinase, DNA-PK.";
RL   Gene 175:271-273(1996).
RN   [11]
RP   CHARACTERIZATION.
RX   MEDLINE=98040126; PubMed=9372844;
RA   Wu X., Lieber M.R.;
RT   "Interaction between DNA-dependent protein kinase and a novel protein,
RT   KIP.";
RL   Mutat. Res. 385:13-20(1997).
RN   [12]
RP   PHOSPHORYLATION OF THR-2609, SER-2612; THR-2638 AND THR-2647.
RX   MEDLINE=22304649; PubMed=12186630;
RA   Douglas P., Sapkota G.P., Morrice N., Yu Y., Goodarzi A.A., Merkle D.,
RA   Meek K., Alessi D.R., Lees-Miller S.P.;
RT   "Identification of in vitro and in vivo phosphorylation sites in the
RT   catalytic subunit of the DNA-dependent protein kinase.";
RL   Biochem. J. 368:243-251(2002).

CC   -!- FUNCTION: SER/THR KINASE INVOLVED IN DNA DOUBLE-STRANDED BREAK
CC   REPAIR, V(D)J RECOMBINATION AND MODULATION OF TRANSCRIPTION. MUST
CC   BE BOUND TO DNA TO EXPRESS ITS CATALYTIC PROPERTIES.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: FORMS A COMPLEX WITH THE HETERODIMER KU70/KU80
CC   AUTOANTIGEN. INTERACTS WITH DNA-PKCS INTERACTING PROTEIN (KIP)
CC   WITH THE REGION UPSTREAM THE KINASE DOMAIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=1;
CC   IsoId=P78527-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=P78527-2; Sequence=VSP_004708;
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   ---

CC   ---
DR   EMBL; U47077; AAB39925.5; -.
DR   EMBL; U34994; AAC50210.3; -.
DR   EMBL; AY316117; AAP69525.1; -.
DR   EMBL; U63630; AAC52019.2; -.
DR   EMBL; U90415; AAB51722.1; -.
DR   EMBL; L27425; AAA79244.1; -.
DR   EMBL; A3052953; BAB79635.1; -.
DR   EMBL; U35835; AAA79184.1; -.
DR   EMBL; AY030284; AAK40350.1; -.
DR   PIR; A57099; A57099.
DR   PIR; G02083; G02083.
DR   SWISS-2DPAGE; P78527; HUMAN.
DR   Genew; HGNC:9413; PRKDC.
DR   MIM; 600899; -.
DR   GO; GO:0004672; F:protein kinase activity; TAS.
DR   GO; GO:0006464; P:protein modification; TAS.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR003152; FATC.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; FALSE_NEG.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Transferase; Serine/threonine-protein kinase; DNA repair;
KW   Nuclear protein; Alternative splicing; Phosphorylation.
FT   DOMAIN   2436   3212   KIP BINDING.
FT   DOMAIN   3748   4128   PI3K/PI4K.
FT   MOD_RES   2609   2609   PHOSPHORYLATION (AUTO-).
FT   MOD_RES   2612   2612   PHOSPHORYLATION (AUTO-).
FT   MOD_RES   2638   2638   PHOSPHORYLATION (AUTO-).
FT   MOD_RES   2647   2647   PHOSPHORYLATION (AUTO-).
FT   VARSPLIC   3799   3829   Missing (in isoform 2).
FT   /FTId=VSP_004708.
FT   CONFLICT   405   405   D -> Y (IN REF. 2; AAC50210).
FT   CONFLICT   1008   1008   A -> S (IN REF. 8; AAC50210).
FT   CONFLICT   3660   3660   N -> T (IN REF. 8).
FT   CONFLICT   3817   3817   L -> W (IN REF. 8).
FT   CONFLICT   3862   3862   A -> P (IN REF. 8).
FT   CONFLICT   4031   4031   I -> V (IN REF. 9).
SQ   SEQUENCE   4128 AA; 469079 MW; AC6E747FEB09F3E5 CRC64;

Query Match   6.0%; Score 564; DB 1; Length 4128;
Best Local Similarity   19.9%; Pred. No. 1.1e-25;

```
Qy   1152  ISLSRGNILQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPOLIARIDTPRP  1211
                :  :   .::    ::          |||:|  :::|   :::
Db   1863  ISEAPGNR-KEMLSKAT-------EDICSHVBEALQHCPTYIWYFVLTQLLSRLLHSHQ   1913

Qy   1212  LVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKLIKNMCEHS---NTLVQQ  1268
                |||:|  ::|  :: :  :::     ::      :::  :::|
Db   1914  SSAQIMHILLSLAVEYPSHILWYITALVNSNSSKRVLRGKHILEKYRQHSQNPHDLVSS   1973

Qy   1269  AMMVSEBLIRVAIlWHHMWHGLEEASRLYFGERNVKGMFEVEPLHAMMERGPQTL-KB   1327
                : ::: ||  :                    ::|        :::|  ::|:
Db   1974  ALDLTKALIRVC-----------------------LQDVKSITSRSGKSLEKD       2003

Qy   1328  TSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVS   1387
                  :  |    ::                 ::  :|    :::|::|  :  :::
Db   2004  FKFDMNVAPSAM-------------------------VVPVRKNLDI--IS          2027

Qy   1388  PKLLMCRDLELAVPGTYDPNQPIRIQSIAPSLQVITSKQRPRKLILMGSNGHEFVFLLK   1447
                |::|        ::  :|||::|:   |::| ||:::::::|:|:|:|:|||||:||||
Db   2028  P-------LESNSMRGYQPFRPVVSIIRFGSSYKVFSSLKKPKQLNIIGSDGNIYGMCK   2080

Qy   1448  GHEDLRQDERVMQLFGLVNTLlANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLH   1507
                |||:|:|:||:||||:||||||||||:::|| |: |||:|::|||:|||:|||:|:|:|
Db   2081  -KEDVRQDNQYMQFATTMDFLLSKDIASRKRSLGINTYSVLSLREDCGILEMVPNVVTLR   2139

Qy   1508  ALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLL--WL   1565
                |::|||:|:||::|:                 :|||||:||:::::::|||:     :|
Db   2140  SILSTKYESLKIKYSLK---------SLHDRKQHTAVDGKLEFYMEQV-----DKFPPILYQWF  2189

Qy   1566  --KSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNI.MLDRLSGKILHIDFGDCFE   1623
                  :|::|  |||::|::|:|:|:|:||||:|:|||||||||:  |:|:|||:|:|::|||:|
Db   2190  LBNFPDPINWFNARNTYARSYAVMAMVGHIILGLGDRHCENILLDIQTGKVLHVDF-DCLF   2248

Qy   1624  VAMTREKFPEKIPFRLIRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVY   1683
                ::||:|||:|||:||:||:::|:|||||||||:|:|::||||:::|:|:|:::|::|:|:|
Db   2249  EKGKRLPVPEIVPFRLTPNLLDALGIICTEGGFFKKSSEVTLALMRKNEVALMNVIETIMY   2308

Qy   1684  DPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESISFI   1743
                :                      :           :|:|:             :   |:|
Db   2309  D-------RNMD--------------------------------------------HSI--  2316

Qy   1744  GDGLVKPEALNKAIQLINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYI   1803
                :||||:|||||||||||:|||:|       ::|:|||||:||::|||||||:||::|:|
Db   2317  -QKALKVL----RNKIRGIDPQDGLVSVAGQTETLIQEA:SEDNLSKMYI           2362

Qy   1804  GWCPFW   1809
           ||||||
Db   2363  GWLPFW   2368
```

RESULT 12

```
PRKD_HUMAN
ID   PRKD_HUMAN   STANDARD;   PRT;   4128 AA.
AC   P78527; P78528; Q13337; Q13327; Q14175; Q72611; Q96SE6; Q9UME3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   DNA-dependent protein kinase catalytic subunit (EC 2.7.1.37) (DNA-
DE   PKcs) (DNPK1).
GN   PRKDC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervical carcinoma;
RX   MEDLINE=95401275; PubMed=7671312;
RA   Hartley K.O., Gell D., Smith G.C.M., Zhang H., Divecha N.,
RA   Connelly M.A., Admon A., Lees-Miller S.P., Anderson C.W.,
RA   Jackson S.P.;
RT   "DNA-dependent protein kinase catalytic subunit: a relative of
RT   phosphatidylinositol 3-kinase and the ataxia telangiectasia gene
RT   product.";
```

```
Qy    157  YGHKVNIGMIDQSRDASAVSLSESKSSQDSSDYS------TSEMLVMNGNLPLDEFYPAV  210
Db   1054  ----NKYVINQNLDDIEVYL-RRKQTERSIDFTPKKVGQTSDITLVLGAL-LDTSHK--  1104

Qy    211  SMVALMRIFRDQSLSHHTMVVQAITFIFKSLGLKCVQFLPQVMPTFLNVIRVCDGAIRE  270
Db   1105  ------------FRN------------LDKDLCEKCAK------CISMIGVLDVTKHE   1132

Qy    271  FLFQQLGMLVSFVKSHIRPYMDEIVTLMREFWVMNTSIQSTIILLIEQIVALGGEFKLY   330
Db   1133  F------KRTTYSBENEVDLNDSVQTIKFLIWVINDI----------------------  1163

Qy    331  LPQLIPHMLRVFMHDNSPGR--IVSIKLLAAIOLFGANLD--DYLHL-LLPPIVLFDAP  385
Db   1164  ---LVP----AFWQSENPSKQLFVALVIQESLKYCGLSSESWDMNHKELYPNEAKLWE--  1214

Qy    386  EAPLPSRKAALETVDRLTESLDFTDYASRIIHPIVRTLDQSPELRSTAMDTLSSLVFQLG  445
Db   1215  ---------------------KFNSVSKTTIYPL--------------LSSL--YLA   1234

Qy    446  KKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGYTILADEEEDPLIYQHRMLRGQGDAL  505
Db   1235  QSWKEYVPL------KYPSNNFKEGYKIW--VKRFTL-------------------DLL  1266

Qy    506  ASGPVETGPMKKLHVSTINLQXAWGAARR--------VSKD--------------DWL-  541
Db   1267  KTGTTENHP--LHVFSSLIREDDGSLSNFLLPYISLDIIIKAEKGTPYADILNGIIIEF  1323

Qy    542  EWLRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSC--WSELNEDQQDELIR  599
Db   1324  DSIFTCNLEGMNNLQVDSLRMCY--------ESIFR-VFEYCKKWATEFKQNYSKLHG   1372

Qy    600  SIELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALFY  659
Db   1373  TFIIKDT----KTTNMLLRIDEFLRTTPSD----------LLAQRSLETDSFERSALY   1416

Qy    660  KELEFQKGP---------TPAILESLISINNKLQQPEAAAGVLE----YAMKHFGELEIQATW  709
Db   1417  LEQCYRQNPHDKNQNGQLLKNLQIITYBEIGDIDSLLDGVLRTFATGNLVSKIEELQYSENW  1476

Qy    710  YEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCE---------K  760
Db   1477  --KLA--QDCFNVLGK-----FSDDPKT--TTRMLKSMYD-HQLYSQIISNSSFHSSDGK  1524

Qy    761  WTLVNDETQAKMARMAAAAAWG-------LGQWDSMEEYTCMIPRDTHDGAFY----RAVLAL  812
Db   1525  ISLSPDVKEWYSIGLEAANLEGNVQTLKNW--VEQIESLRNIDDREVLLQYNIAKALLIAI  1582

Qy    813  HQDLFSLAQQCIDKARDLL------DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQY  865
Db   1583  SNEDPLRTQKYIHNSPFLIGTNFITSSKETTLLKKQNLLMKLHSLYDLSFLSSAKDKPEY  1642

Qy    866  K----LVPERREIIRQIWERLQGCQRIVEDW----QKILMVR----SLVVSPHEDM---R  911
Db   1643  -------------RIGADFVPNHYIILSMRKSFDQLKKVNEQADADLGK   1688

Qy    912  TWLKYASLCGKSGRLALAHKTLVLLIGVPDSRQLDHPLPTVHPQVTYAYMKNVMKSARKI  971
Db   1689  TFFTLAQLCARNNARLDIASEBSLM-------HCLERRLPQAELEFABILWKQGEND   1736

Qy    972  DAFQHMQHFVQTMQQQAQHAIATEDDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPK  1031
Db   1737  RALKLVQBIHEKYQENS--SVNARDR--------AAVLLKFTEW-LDLS--NNSASEQ   1781

Qy   1032  VLQYYSAATEHDRSWYKAWHAWAVMNFRAVLHYKHQNQARDEKKKLRHASGANITNATTA  1091
Db   1782  IIKQYQDIFQIDSKWDKPYYSIG------LYY--SRILERKKAEGYITNGRPEYRAIS  1831

Qy   1092  ATTATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLLSKTLLMYTVPAVQGFFRS  1151
Db   1832  YFLLAAFKNTA---------------------KVRENLPKVITFWLDIAA----AS   1862
```

```
Db   2368 MQTYLEKAVKVAGSYDGNSRELRNGQMKAFLSLARFSDTQYQRIENYMKSSEFENKQTLL 2427
                                                                          repair and meiotic recombination in Saccharomyces cerevisiae.";
                                                                       RT
Qy   1078 RHA-SGANITNATTAATTAATATTTASTEGSNSESSEABSTENSPTPSPLQKKVTEDLSKT 1136   RL Nucleic Acids Res. 22:3104-3112(1994).
               :|:          :    :   :  :         ||:  :                  RN [2]
Db   2428 KRAKEEVGLLREHKIQTNRYIVKV-------QRELELDECA------LRALREDRKR-    2471   RP SEQUENCE FROM N.A.
                                                                          RA Weinert T.A., Harlow D.;
                                                                          RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
Qy   1137 LLMYTVPAVQGFPRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWL 1196   RN [3]
               :| ||::::|:|::: :|| :      :                  ::|        RP SEQUENCE FROM N.A.
Db   2472 --FLCKAVENYINCLLSGEEHDLW-VFRLCSLWLENSGVSEVNGMMKKDCMKISSYKFL  2527   RC STRAIN=S288c;
                                                                          RX MEDLINE=94378717; PubMed=8091856;
Qy   1197 QVIPLIARIDTPRPLVGRL-----IHQLLTDIGRYHPQALIYPLIV----------ASKS 1242   RA Becam A.-M., Cullin C., Grzybowska E., Lacroute F., Nasr F.,
               :|:      | ||   |       |:|  :|:| :|   :|:|  :            RA Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
Db   2528 PLMYQLAARMGT--KMTGGLGFHEVLNNLISRISLDHPHTLPIIIALANANKDEFLSKP  2585   RA Herbert C.J.;
                                                                          RT "The sequence of 29.7 kb from the right arm of chromosome II reveals
Qy   1243 TTTARHNAANKILKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEBEASRLYFCER 1302   RT 13 complete open reading frames, of which ten correspond to new
                       ::| :    : :|:|: :  :: ||:                         RT genes.";
Db   2586 ETTRR----SRITKST8KENSHLDE-----------------DRTEAATRIIHSIR     2620   RL Yeast 10:S1-S11(1994).
                                                                          CC -!- FUNCTION: Required for cell growth, DNA repair and meiotic
Qy   1303 N----VKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDL  1357   CC     recombination.
               |:      :|: :|        : : :|: :  ::  |               :    CC -!- DEVELOPMENTAL STAGE: Induced during meiosis.
Db   2621 SKRCKMVKDM-EALCDAYIILAN----------------MDASQW-RAQRKGINIP--   2658   CC -!- SIMILARITY: IN THE N-TERMINAL REGION, TO S.POMBE CUT1.
                                                                          CC -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
Qy   1358 TQAWDLYYHVFRRISKQLPOLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIA 1417   CC ---------------------------------------------------------------
               ::|:|:|||: | |     :|:|::|                               CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
Db   2659 ---------ANQPITKLKNLE-DVVVFTM----EIKVDPTGEY--ENLVTIKSFK      2697   CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                          CC  the European Bioinformatics Institute.  There are no restrictions on its
Qy   1418 PSLQVITSKQRPRKLITLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLR 1477   CC  use by non-profit institutions as long as its content is in no way
               ||:| :  |: :: :|||: :|||| |||| :|||:|   |::|: :::|        CC  modified and this statement is not removed. Usage by and for commercial
Db   2698 TEFRLAGGLNLPKIIDCVGSDGKERRQLVKGRDDLRQDAVMQQVFQMCNTLLQRNTETRK 2757   CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
                                                                          CC  or send an email to license@isb-sib.ch).
Qy   1478 KNLSIQRYAVIPLSTNSGLIGW-------------------VPHCDTLHALIRDYR----- 1519   CC ---------------------------------------------------------------
               :|| | :|  |:|:||:||                    :   |:|:          DR EMBL; X75891; CAA53494.1; -.
Db   2758 RKLTICTYKVVPLSQRSGVLEWCTGTVPIGEYLVNSEDGAH--RRYRPNDFSANCCQKK  2814   DR EMBL; U31109; AAA74482.1; -.
                                                                          DR EMBL; Z36005; CAA85094.1; -.
Qy   1520 LLNIEHR-MLRMAPDYD-HLTLMQKVE-VFBHAVNNTAGDDLAXLLWLKSPSSVWFDRR  1577   DR EMBL; D11088; BAA01860.1; -.
               :::       :| |    ::| ::|      :      |:::|| :            DR PIR; S46005, S46005.
Db   2815 MMEVQKK--SPEEKYDTFMTICQNFEPVFRYFC------MEKFL-----DPAVWFSKR   2859   DR GermOnline; 138679; -.
                                                                          DR SGD; S0000340; MEC1.
Qy   1578 TNYITRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPF 1637   DR GO; GO:0005634; C:nucleus; IC.
               ||: ||| |:||||| ||:|||||:|:|||||:|::||||||| |  | ::| |    DR GO; GO:0000076; P:DNA replication checkpoint; IMP.
Db   2860 LAYTRSVATSSTVGYILGLGDRHVQNLILNEQSARLVHIDLGVAFBQGKILPT-PETVPF 2918   DR GO; GO:0007131; P:meiotic recombination; IMP.
                                                                          DR GO; GO:0006139; P:nucleobase, nucleoside, nucleotide and nucl  . .; IGI.
Qy   1638 RLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVTDPLLNWRL------- 1691   DR InterPro; IPR008938; ARM.
               ||:|: |    || | ::|  |: :|:||  :| :::|     ::|| |         DR InterPro; IPR003151; FAT.
Db   2919 RLSRDIVDGMGITGVEGVPRRCCEKTMEVNRSSQETLLTIVEVLLYDPLFDWTMNPLKAL 2978   DR InterPro; IPR003152; FATC.
                                                                          DR InterPro; IPR000403; PI3_PI4-kinase.
Qy   1692 ---MDTNIKGNKRSRTRTDSYSAGQSVBILDGVELGEPAHKKTGTTVPESIHSFIGDGLV 1748   DR InterPro; IPR008941; TPR-like.
               | | |:|  || :| |:|:   ::| || ::                          DR Pfam; PF02259; FAT; 1.
Db   2979 YLQQRPEDESDLHSTPNADDQECKQSLSDTD---------------------------- 3009   DR Pfam; PF02260; FATC; 1.
                                                                          DR Pfam; PF00454; PI3_PI4_kinase; 1.
Qy   1749 KPEALNKKAIQIINRVRDKJTGRDFSHDDTLDVFTQVELLIKQATSHENLCQCYIGW  1805   DR SMART; SM00146; PI3Kc; 1.
               ||: |:|   :||: ::| ::                                      DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
Db   3010 --QSFNKVAERVLMRLQEKLKG--VEBGTVLSVSGVGQVNLLIQQAMDPKNLSRLFPGW 3062   DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
                                                                          DR PROSITE; PS00290; PI3_4_KINASE_3; 1.
                                                                          KW Meiosis; Transferase; Kinase; DNA repair.
                                                                          FT DOMAIN        2082   2368       PI3K/PI4K.
                                                                          FT CONFLICT       197    197       N -> D (IN REF. 1).
                                                                          FT CONFLICT       716    716       S -> P (IN REF. 2).
                                                                          FT CONFLICT      1255   1255       K -> Q (IN REF. 2).
                                                                          FT CONFLICT      1276   1276       L -> G (IN REF. 2).
                                                                          SQ SEQUENCE   2368 AA;  273339 MW;  C06AEF9F0484A615 CR064;
```

RESULT 11
ESR1_YEAST
ID   ESR1_YEAST   STANDARD;   PRT;   2368 AA.
AC   P38111; Q02580;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   ESR1 protein.
GN   ESR1 OR MEC1 OR SAD3 OR YBR136W OR YBR1012.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94344772; PubMed=8065923;
RA   Kato R., Ogawa H.;
RT   "An essential gene, ESR1, is required for mitotic cell growth, DNA

```
Query Match             6.7%;  Score 628; DB 1;  Length 2368;
Best Local Similarity  20.1%;  Pred. No. 5.1e-30;
Matches  375;  Conservative 305;  Mismatches 614;  Indels 572;  Gaps  77;

Qy    54 VINNVLATIGELAQVSGLEMRKWVD------ELFPIIMDML----QDSSLLAKRQVALWT 103
            :|::|:    |||: :| |: ||        ||: :|:|        ::|  :  : |||
Db   965 VIDSLIAFI--------------------------------LQKWSEFNGKLRNIVYSILDTLIKBKSDLLILKLKPYTTLAL 1014

Qy   104 LGQ---LVASTGY--VVEPYRKYPTLLEVLLNFLKTBQNQGTRREAIRVLGLLGALGP 156
            :|:     ::|  |:    :  || :||::|              :|: ||:|::|
Db  1015 VGKPELGILARDGQFARMVNKLRSTTDLLPIFANNLKSS--------------------- 1053
```

```
CC  -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SKELETAL MUSCLE, TESTIS,
CC      FOLLOWED BY SPLEEN, LUNG, KIDNEY, HEART, LIVER AND THYMUS.
CC      UBIQUITOUSLY EXPRESSED IN EMBRYONAL TISSUES.
CC  -!- DEVELOPMENTAL STAGE: HIGHEST EXPRESSION IN EMBRYONIC CENTRAL
CC      NERVOUS SYSTEM, FROM E13.5 DAY AND DURING THE WHOLE CEREBELLAR
CC      DEVELOPMENT. DECREASED EXPRESSION WHEN MATURATION OCCURS.
CC  -!- PTM: Phosphorylated (By similarity).
CC  -!- DISEASE: ATM-DEFICIENT MICE SHOW A PHENOTYPE SIMILAR TO HUMAN
CC      ATAXIA TELANGIECTASIA (AT) AND CONSISTENTLY DEVELOP IMMATURE T-
CC      CELLS MALIGNANCIES.
CC  -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
DR  EMBL; U43678; AAC52673.1; -.
DR  MGD; MGI:1072202; Atm.
DR  GO; GO:0008630; P:DNA damage response, signal transduction re. . .; IMP.
DR  InterPro; IPR003151; FAT.
DR  InterPro; IPR003152; FATC.
DR  InterPro; IPR000403; PI3_PI4_kinase.
DR  Pfam; PF02259; FAT; 1.
DR  Pfam; PF02260; FATC; 1.
DR  Pfam; PF00454; PI3_PI4_kinase; 1.
DR  SMART; SM00146; PI3Kc; 1.
DR  PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR  PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR  PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW  Transferase; Kinase; Nuclear protein; Cell cycle; DNA repair;
KW  DNA-binding; Phosphorylation; Anti-oncogene.
FT  DOMAIN     1380   1389     C-ABL-BINDING (BY SIMILARITY).
FT  DOMAIN     1972   2576     FAT.
FT  DOMAIN     2722   3066     PI3K/PI4K.
SQ  SEQUENCE   3066 AA;  349484 MW;  1EB21EB61405577A CRC64;
```

```
Query Match          6.7%;  Score 628.5;  DB 1;  Length 3066;
Best Local Similarity 18.4%;  Pred. No. 7.2e-30;
Matches 430;  Conservative 356;  Mismatches 742;  Indels 809;  Gaps  97;
```

```
Qy      32 MEPILKALILKLKDPDPDPNPGVINNVLATIGELAQVS-GLEMRKWVDELFIIIMDMLQD 90
            :: :            ||          |     |::::   || :|||  ||| :|||
Db     972 LQPL--SLVCSLHRRDQDVCKTILSNVLHIVTNLGQGSVDMESTRIAQGHFLTVMGAFWH 1029

Qy      91 SSLLAKRQVALWTLGQLVA-------STGYVVEPYRKYPTL-------LEVLLNPLKTEQNQ 138
            ::                      :::|                  ||  :   |:
Db    1030 --LTKEKKCVFSVRMALVKCLQTLLEADPYSEWAILNVKGQDPPVNEAFSQFLADDHHQ 1086

Qy     139 GTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSESKSSQDSDYSTSE----- 193
            |: :::      | |:::|                      |  ::|:|:|
Db    1087 --VRMLA-AGSVNRLFQDMRQG--DFSRSLKALPLKPQQTSFNNA-YTTAEAGIRG 1136

Qy     194 MLVNMGNLP-LDEFYPAVSMV---------------ALMRIFRDQSLSHHHTMVV 232
            ||:||||:|||| :|:            ||       |  |  :|  |: ||
Db    1137 LLCDSQNPDLLDELYNRKSVILMMIAVVLHCSPVCEKQALFALCKSVKENRLEPH--LVK 1194

Qy     233 QAITFIFKSLGLK----CVQFLPQVMPTFLNVIRVCDGAIREFLFQQLG--MLVSFVKS 285
            ::: |||||| ||   | :   |::|:   : : |::| |    |        ::  |:
Db    1195 KVLEKVSESFGCRSLEDFMISHLDYLVLEWIN-LQDTEYSLSSPFMLNYTSIEDFYRS 1253

Qy     286 ------H--IRPYMDEIVTL--MREFWVMNTSIQSTIILIEQIVVALGGEFKLYLPQ 333
            :       ::   ::||::  | | :::|:|::            |:| | ||: ||:
Db    1254 CYKLLIPHLVIRSHFDEVKSIANQIQKCWKS--------LLVD----------CFPK 1292

Qy     334 LIPHMLRVFMDNSPGRIVSIKLLAAIQLF--------GANLDDYLHLLLPPIV---- 379
            : :||:|        : ::|||  : |         || :|:|:||||| |:
Db    1293 ILVHILPFYAYEGIRDSYVSQKRETATKVDTLKGEDFLGKQIDQVFISNLPEIVVELLM 1352

Qy     380 -----------KLFDAPEAP-LPSR-------------------------------KAAL 396
            :          :  ::  ::   ::     ::
```

Alignment region (continued):

```
Db    1353 TLHETADSADSDASQSATALCDFSGDLDPAPNPPYFPSHVIQATFAYISNCHKTKPKSII 1412
Qy     397 ETVDRLTESLD------------------------------------------ 407
            |:                                   :: :
Db    1413 EILSKIPDSYQKILLAICBQAAETNVFKKHRILKIYHLFVSLLLKDIQSGLGGAWAFVL 1472
Qy     408 --------FTDYA-----------------SRI-----IHPIVRT-- 422
            ||| :                :::|      :::|:|:|
Db    1473 RDVIYILIHYINKRSSHFTDVSLRSFSLCCDLLSRVCHTAVTQCKDALESHLHVLVGTLI 1532
Qy     423 -LDQSPELRSTAMDTLSSLVF----------QLGKKYQIFIPMVNKVLVR- 461
            :  : :|::|:: ::::  |           |:
Db    1533 PLVDYQEVQEQVLDLLKYLIVIDNKDKNLSVTIKLLDPFPDHVIFKDLRLTQQKIKYSGG 1592
Qy     462 -----HRINH---QRYDVLICRIVKGY------------------TLADEEDP------- 489
            : ::|            |: : |::         ::::
Db    1593 PFSLLEINHPLSVSAYNPLITRLEGLKDLRRQLEQHKDQMLDLRASQDNPQDGIVVK 1652
Qy     490 ---LIYQHRML--RSGGQDALAS----------GPVETGPMKKLHVSTINLQKAWGAAR- 533
            ::|:|::    :: :: :          ||:
Db    1653 LVVSLLQLSKMAVNQTGEREVLEAVGRCLGEIGPLDPSTIAVQHNKDVSYTKAYGLPEDR 1712
Qy     534 ------------------------------------------------------- 533
Db    1713 ELQWTLIMLTALNNTLVEDSVKIRSAAAATCLKNILATKIGHIFWENYKTSADPMLTYLQP 1772
Qy     534 -------------RVSKDDWLE-------------------------WLRRLS----LEL 551
            :           :: :::                       |:|::     |:|
Db    1773 FRTSRKKFLEVPRSVKEDVLEGLDAVNLWVPQSESHDIWIKTLTCAFLDSGGINSPILQL 1832
Qy     552 LKDSSSPSLRSCW----------ALAQAYNPMARDLFNA----AFVSCW---------SE 588
            |:|::::|::|          |  |  ::|:|:|         ::|:|         |:
Db    1833 LKPMCEVKTDFCQMLLPYLIHDVLLQDTHESWRTLLSAHVRGFFTSCFKHSSQAGRSATP 1892
Qy     589 LNEDQQDELIRSIELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVL----- 642
            |  :| ::| :|::   ::| :       :   :::  :: : |
Db    1893 ANSDSBSENFLRCCLDKKSQ-------------RTMLAVVDYLRRQKRPSSGTAPDDAFWLDLNYLE 1946
Qy     643 LGERAAKCRAYAKALHYKEL------------------------------EFQKGPTAILES 675
            |::    |  |: |:::                              |  : | : ::
Db    1947 VAKVAQSCSAHFTALLYAEIYSDKKSTDEQKKRSPTFEEGSQGTTISSLSEKSKEETGIS 2006
Qy     676 ----LISINNKLQQPEA----AAGVLEYAMKGHFGELEIQATWYEKLHEWEDALVAYDKRM 727
            :     : :| ::||       | |: : : | |:| : || : ||::|| |:|:|
Db    2007 LQDLLLEIYRSIGEPDSLYGCGGKMLQPLTRIRTYBHEAT------WEKALVTYD--L 2057
Qy     728 DTNKDDPELMLGRMRCLEALG-----------------EWGQLHQQCCEKWTLVNDETQA 770
            :::|:   :|:| :  :|                   :|| | :    |:|| :::|
Db    2058 ETSISSSTRQSGIIQALQNLGLSHILSVLKGLDYERREW----------C------A 2099
Qy     771 KVARMAAAAAWGLGQWDSMERYTCMIPRDIHDGAFYRAVLALHQDLF----SLAQQCID 825
            |    :|          ::: :|::| :|  ||  : :  :|:|:    :|::|:|
Db    2100 ELQELRYQAAWRNMQWGLCASAGQEVEGTSYHESLIYNALQCLRNREFSTFYESLRYASLF 2159
Qy     826 KARDLLDAELTAMAGESYSRAYGAMVSCHMLSELE----EVIQYKLVPERRRIIRQIWWER 882
            |  :|::| :|:|||| | :|  | :: |:  |:   ||: |:|  ||:|| |::| |:
Db    2160 RVKEV--EELSKGSLESVYSLYPTLSRLQAIGELENSGELFSRSVTDRERS---EAYWK- 2213
Qy     883 LQGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPS 942
            |::::|||:|::| :::| | ||  :|:   |::    |: :: : :|:::|  || ::
Db    2214 -----------WQK------HSQL---LKDSDFSFQEPLMALRTVILETLVQKEME 2249
Qy     943 RQ--LDHPLPTVHPQVTYAYMKNMWKSAR-----KIDAFQ------EMQHFVQT 983
            ::   | :|   |:  : | |:  | |:     |:|       ::::| |:
Db    2250 RSQGACSKDILTKH-LVEFSVLARTFKNTLQPERAIFKIKQYNSAICGISEWHLEE-AQV 2307
Qy     984 MQQQAQHAIAT-------------------EDQQHKQELHKLMARCFLKLGEW--------QLN 1020
            ::::|: |:|                 :|::| ||:: |: | | |          |::
Db    2308 FWAKKEQSLALSILKQMIKKLDSSFKDKSNDAGLKVIYABCLRVCGSWLAETCLENPAVI 2367
Qy    1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVNN---FEAVLHYKHQNQARDEKKKL 1077
            |::::| |||  |:| ||  |:::  : || | ::::    :: :::|::|::: :::
```

```
Qy    749 ------------------------EWGQLHQCCCEKWTLVNDETQAKMARMAAAAAWGLGQWDSM 789
                                  ||:|:|                        ||||  ||
Db   2075 HILSVYLKGLDYENKDWCPELEELHYQ------------------------AAWRNMQWDHC 2112

Qy    790 EEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAGESYRAYGA 849
              :::|  |:::: |  |:: |   |      |:|:   |: |  ||     |: :|
Db   2113 TSVSKEVEGTSYHESLYNALQSLRDREFSTFYESLKYARVKEVEEMCKRSLESVYSLYPT 2172

Qy    850 MVSCHMLSELEEV---------IQYKLVPE-------------------RREIIRQ 877
              : ||| :|          :|||                     :|||||:
Db   2173 LSRLQAIGELESIGELFSRSVTHRQLSEVYIKWQKHSQLLKDSDFSFQEPIMALRTVILE 2232

Qy    878 IWWER-LQSCQR--IVEDWQKILMVRSLVVS------PHEDMRTWLKYASL-CGKS-- 923
          |  |:  | :|::||:  ::|::                || :  |: | :  |
Db   2233 ILMEKEMDNSQRECIKDILTKHLVELSILARTFKNTQLPERAIFQIKQYNSVSCGVSEWQ 2292

Qy    924 --GRLAiAHKTIVLLLGV-DPSRQLDHPLPTVHPQVTYAYMKNM-----WKSARKIDA 973
            ::: |:|| |:::|:|  :| : | |:  :: |:|     |||: :| :
Db   2293 LEEAQVFWAKKEQSLALSIIKQMIKKLDASCAANNPSLKITYTECLRVCGNWLAE---TC 2349

Qy    974 FQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCPLKLGEWQLNLQINESTIPKVL 1033
             |  |   |      ||   : ||:: ||||||: || | |  : | | :   :|:
Db   2350 LENPAVIMQTYLEKAVEVAGNYDGESSDELRNGKMKAFLSLARF------SDTQYQRI 2401

Qy   1034 QYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAAT 1093
          :: :|| :|:                                       :|  |  |::|
Db   2402 ENYMKSSEFE----------------NKQALL-----KRAKBEVGLLREHKIQ-TNRYTVKV 2441

Qy   1094 TAATATTTASTEGSNSESEARESTENSPTPSELQKKVTEDLSKTLLMYTVPAVCGFFRSIS 1153
          |  || :     ||:   |||: :|    : | | :    |: | : :|:  : : ::
Db   2442 ---QRELELDELA------LRALKEDRKR---FLCKAVENYINCL- 2474

Qy   1154 LSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLV 1213
          ||:  ||  |: |:  ||  |   ::|    ::|||   ||| ::  |:  | |  |:|||
Db   2475 LSGEEHDMWVFRLCSWLENSGVSEVNGMKKRDGMKIPTYKFLPLMYQLAARMGT-KMM 2532

Qy   1214 GRL-----IHQLLTDIGRYHPQ---ALIYPLTVASKSTTTARHNAA--NKILKNMCEHSN 1263
          || |       :|: : ||       :|| | | ||:: | |:|   :|:|  | : |
Db   2533 GGLGFHBEVLNNLISRISMDPHHTLFLILALANANRDEFLTKPEVARRSRITKNVPKQSS 2592

Qy   1264 TLVQQAMMVVSEELIRVAILWHEMWHEGLEBASRLYFGGERNVKGMFPVLEPLHAMMERGPQ 1323
          || |  : ::|  :|          ||                      || : |:| ||
Db   2593 QLDE---------------------DRTERAANRI----------ICTIRSRRPQ 2615

Qy   1324 TLKETSFNQAYGRDLMEACEWCRKYMKSGNVKDLTQAWDLYHVFRRISKQLP-------Q 1377
          | ::|::  || | |:  :  | |:  |::||:|  | |::  :| :| :|
Db   2616 MVR--------SVEALCDAYIILANL-DATQ-WKT----QRKGINIPADQPITK 2655

Qy   1378 LTSLELQYVSPKLLMCRDLELAVPGTVDPNQPIIRIQSIAPSLQVITSKQRPKLITLMGS 1437
          | || |  ||:|| | : |  ||  |:  |     ::  | :|:|| ||:|| | : ||
Db   2656 LKNLE-DVVVPTM----EIKVDHTGEYG---NLVTIQSFKARFRLAGGVNLPKILDCVGS 2707

Qy   1438 NGHEFVFLLKGHEDILRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLI 1497
          ||:|| ||: |||:|:|  |  | :|| :| |  |||: |::| |||:|:    ::| ||
Db   2708 DGKERRQLVKGRDDLRQDAVMQQVFQMCNTLLQRNTETRKRKLTICTYKVVPLSQRSGVL 2767

Qy   1498 GW------VPHCDTLHALIRDYREKKKIILNIEBHRIMLRMAPDYDHLTLMQKVEV 1546
          ||        ||  |: |: |:::||| : ||  || ||  ||||:  |:|: |
Db   2768 EWCTGTVPIGEFLVNNEDGAH---KRYRPNDFSAFQCQKKMM-----EVQKKSFEEKYEV 2819

Qy   1547 FEHAVNNTAGDDLAKLLWL-KSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLM 1605
          |   |:| ||||||:||   ||| ||:      |: |||   | | |:||||||  | |
Db   2820 FMDVCQNF--QPVFRIFCMEKFLDPAIWFEKRLAYTRSVATSSIVGYIILGLGDRHVQNIL 2877

Qy   1606 LDRLSGKILHIHFDGDCFEVAMTREKFPEKIFPRLLRMLTNAMEVTGLDGNYRITCHTVME 1665
          |:|:| ||:| ||  |  | |     | |::|  | | | |   :|| ||| | :| |::
Db   2878 INEQSAELVHIDLGVAFEQGKILPT-PETVPFRLTRDIVDGMGITGVEGVFRRCCEKTME 2936

Qy   1666 VLREHKDSVMAVLEAPVIDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVBLGE 1725
          |:    || | ||::   ||| | :   :|  |  | :::    :| : :|| | :
Db   2937 VMRNSQETLLTIVEVLLYDPLFDWTM---NPLKALYLQQRPE------DETELEPTLNADD 2988

Qy   1726 PAIHKTGTTVPESIHSFIGDGLVKPEALNKKAIQIINRVRDKLTGRDSHDDTLDVPTQV 1785
```

```
Db   2989 QECKRNLSDIDQS------------------------------FDKVAERVLMRLQEKLKG--VEEGTVLSVGGQV 3032
                ||:|::| : :|                            ||:|  :: | |::|||:   :||| || |||
Qy   1786 ELLIKQATSHENLCQCYIGW 1805

Db   3033 NLLIQQAIDPKNLSRLFPGW 3052
```

RESULT 10
```
ATM_MOUSE
ID   ATM_MOUSE    STANDARD;    PRT;    3066 AA.
AC   Q62388;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Serine-protein kinase ATM (EC 2.7.1.37) (Ataxia telangiectasia mutated
DE   homolog) (A-T, mutated homolog).
GN   ATM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96299738; PubMed=8661102;
RA   Pecker I., Avraham K.B., Gilbert D.J., Savitsky K., Rotman G.,
RA   Harnik R., Fukao T., Schroeck E., Hirotsume S., Tagle D.A.,
RA   Collins F.S., Wynshaw-Boris A., Ried T., Copeland N.G., Jenkins N.A.,
RA   Shiloh Y., Ziv Y.;
RT   "Identification and chromosomal localization of Atm, the mouse homolog
RT   of the ataxia-telangiectasia gene.";
RL   Genomics 35:39-45(1996).
RN   [2]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=97152562; PubMed=9000145;
RA   Lakin N.D., Weber P., Stankovic T., Rottinghaus S.T., Taylor A.M.R.,
RA   Jackson S.P.;
RT   "Analysis of the ATM protein in wild-type and ataxia telangiectasia
RT   cells.";
RL   Oncogene 13:2707-2716(1996).
RN   [3]
RP   DEVELOPMENTAL STAGE.
RC   STRAIN=B6C3-F1;
RX   MEDLINE=98362438; PubMed=9697112;
RA   Soares H.D., Morgan J.I., McKinnon P.J.;
RT   "Atm expression patterns suggest a contribution from the peripheral
RT   nervous system to the phenotype of ataxia-telangiectasia.";
RL   Neuroscience 86:1045-1054(1998).
RN   [4]
RP   BETA-ADAPTIN BINDING.
RX   MEDLINE=98374320; PubMed=9707615;
RA   Lim D.-S., Kirsch D.G., Canman C.E., Ahn J.-H., Ziv Y., Newman L.S.,
RA   Darnell R.B., Shiloh Y., Kastan M.B.;
RT   "ATM binds to beta-adaptin in cytoplasmic vesicles.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:10146-10151(1998).
CC   -!- FUNCTION: Involved in signal transduction, cell cycle control and
CC       DNA repair. May function as a tumor suppressor. Necessary for
CC       activation of ABL1 and SAPK. Phosphorylates p53, NFKBIA, BRCA1,
CC       CTIP, NIBRIN (NBS1), TERF1, and RAD9. May play a role in vesicle
CC       and/or protein transport. Inhibited by wortmannin. Could play a
CC       role in T-cell development, gonad and neurological function (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: Exists in monomeric and tetrameric state. Binds DNA ends.
CC       P53, ABL1, BRCA1, NIBRIN (NBS1) and TERF1. Part of the BRCA1-
CC       associated genome surveillance complex (BASC), which contains
CC       BRCA1, MSH2, MSH6, MLH1, ATM, BLM, PMS2 and the RAD50-MRE11-NBS1
CC       protein complex. This association could be a dynamic process
CC       changing throughout the cell cycle and within subnuclear domains
CC       (By similarity).
CC   -!- SUBCELLULAR LOCATION: PRIMARILY NUCLEAR. FOUND ALSO IN ENDOCYTIC
CC       VESICLES IN ASSOCIATION WITH BETA-ADAPTIN (BY SIMILARITY).
```

RT   damage.";
RL   Nature 387:520-523 (1997).
RN   [14]
RP   P53 BINDING, AND KINASE ACTIVITY.
RX   MEDLINE=99057351; PubMed=9843217;
RA   Khanna K.K., Keating K.E., Kozlov S., Scott S., Gatei M., Hobson K.,
RA   Taya Y., Gabrielli B., Chan D., Lees-Miller S.P., Lavin M.F.;
RT   "ATM associates with and phosphorylates p53: mapping the region of
RT   interaction.";
RL   Nat. Genet. 20:398-400(1998).
RN   [15]
RP   BETA-ADAPTIN BINDING.
RX   MEDLINE=98374320; PubMed=9707615;
RA   Lim D.-S., Kirsch D.G., Canman C.E., Ahn J.-H., Ziv Y., Newman L.S.,
RA   Darnell R.B., Shiloh Y., Kastan M.B.;
RT   "ATM binds to beta-adaptin in cytoplasmic vesicles.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:10146-10151(1998).
RN   [16]
RP   PHOSPHORYLATION OF P53.
RX   MEDLINE=98404273; PubMed=9733514;
RA   Banin S., Moyal L., Shieh S.-Y., Taya Y., Anderson C.W., Chessa L.,
RA   Smorodinsky N.I., Prives C., Reiss Y., Shiloh Y., Ziv Y.;
RT   "Enhanced phosphorylation of p53 by ATM in response to DNA damage.";
RL   Science 281:1674-1677(1998).
RN   [17]
RP   PHOSPHORYLATION OF P53, AND MUTAGENESIS OF ASP-2870 AND ASN-2875.
RX   MEDLINE=98404274; PubMed=9733515;
RA   Canman C.E., Lim D.-S., Cimprich K.A., Taya Y., Tamai K.,
RA   Sakaguchi K., Appella E., Kastan M.B., Siliciano J.D.;
RT   "Activation of the ATM kinase by ionizing radiation and
RT   phosphorylation of p53.";
RL   Science 281:1677-1679(1998).
RN   [18]
RP   DNA BINDING.
RX   MEDLINE=99432198; PubMed=10500142;
RA   Smith G.C.M., Cary R.B., Lakin N.D., Hann B.C., Teo S.-H., Chen D.J.,
RA   Jackson S.P.;
RT   "Purification and DNA binding properties of the ataxia-telangiectasia
RT   gene product ATM.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:11134-11139(1999).
RN   [19]
RP   PHOSPHORYLATION OF BRCA1.
RX   MEDLINE=20018333; PubMed=10550055;
RA   Cortez D., Wang Y., Qin J., Elledge S.J.;
RT   "Requirement of ATM-dependent phosphorylation of brca1 in the DNA
RT   damage response to double-strand breaks.";
RL   Science 286:1162-1166(1999).
RN   [20]
RP   IDENTIFICATION OF ATM AS MEMBER OF BASC.
RX   MEDLINE=20245492; PubMed=10783165;
RA   Wang Y., Cortez D., Yazdi P., Neff N., Elledge S.J., Qin J.;
RT   "BASC, a super complex of BRCA1-associated proteins involved in the
RT   recognition and repair of aberrant DNA structures.";
RL   Genes Dev. 14:927-939(2000).
RN   [21]
RP   PHOSPHORYLATION OF NIBRIN.
RX   MEDLINE=20227312; PubMed=10766245;
RA   Lim D.-S., Kim S.-T., Xu B., Maser R.S., Lin J., Petrini J.H.J.,
RA   Kastan M.B.;
RT   "ATM phosphorylates p95/nbs1 in an S-phase checkpoint pathway.";
RL   Nature 404:613-617(2000).
RN   [22]
RP   PHOSPHORYLATION OF NIBRIN.
RX   MEDLINE=20296355; PubMed=10839545;
RA   Wu X., Ranganathan V., Weisman D.S., Heine W.F., Ciccone D.N.,
RA   O'Neill T.B., Crick K.E., Pierce K.A., Lane W.S., Rathbun G.,
RA   Livingston D.M., Weaver D.T.;
RT   "ATM phosphorylation of Nijmegen breakage syndrome protein is required
RT   in a DNA damage response.";
RL   Nature 405:477-482(2000).
RN   [23]
RP   PHOSPHORYLATION OF NIBRIN.
RX   MEDLINE=20264381; PubMed=10802669;

RA   Gatei M., Young D., Cerosaletti K.M., Desai-Mehta A., Spring K.,
RA   Kozlov S., Lavin M.F., Gatti R.A., Concannon P., Khanna K.K.;
RT   "ATM-dependent phosphorylation of nibrin in response to radiation
RT   exposure.";
RL   Nat. Genet. 25:115-119(2000).
RN   [24]
RP   PHOSPHORYLATION OF CTIP.
RX   MEDLINE=20365735; PubMed=10910365;

```
Query Match             6.7%;  Score 631;  DB 1;  Length 3056;
Best Local Similarity   20.3%;  Pred. No. 5e-30;
Matches 418;  Conservative 323;  Mismatches 787;  Indels 532;  Gaps  87;

Qy         16 MLGHLVSNAPRLIRPYMEPILKALLIKL----KDPDPDNPGVINNVLATIG-ELAQYSG   70
              :: ||||                     |     |     | |||::|| |::|:|||
Db       1255 LIPHLV----IRSHFDEV-KSIANQIQEDWKSLLTDCFPKILVNLPYFAYEGTRDSG 1307

Qy         71 LEMRKKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRKYPTLLEVLLN  130
              |  :     |:  |||| :     ||:  :|   |                  ||:|: |:
Db       1308 MAQQR--ETATKVYDMLKSENLLGKQ-------IDHLFIS---------NLPEIVVE 1346

Qy        131 FLKT-EQNGTRREAIRVLGLLGALDPYKH--------KVNIGMIDQSRDASAVS---  176
              |: |  :       |  : :|| ||   |:          |   ||  |  :| |
Db       1347 LLMTLHEPANSSASQSTDLCDFSGDLDPAPNPHFPSHVIKATFAYISNCHKTKIKSIIE 1406

Qy        177 -LSESKSSQDSDYSTSEMLVNMGNL----------PLDEYPAVSMV--------ALMRIFR  220
              | ||||                |    |          |    |:|:          | |||
Db       1407 IISKSPDSYQKILLAICEQAAETNNVYKKHRILKIYHLPVSLLLKDIKSGLGGAWAFVLR 1466

Qy        221 D--QSLSHHTMVVQAITFI-FKSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLG  277
              | |  :|: :  :| :|| ::| ::  :| | :|: | ||:||:|       :|
Db       1467 DVIYTLIHYINQRPSCIMDVSLRSFSL-CCDLLSQVCQT--AVTYCKDALENHLHVIVG 1522

Qy        278 MLVSFVKSHIRPYMDEIVTLMREFWV--MNTSIQSTILL------------IEQIV  320
              |: |        ||  ::  |  |  |      ||  |||            | ||
Db       1523 TLIPLVYEQVE-VQKQVLDLIKYLVIDNKDNENLYTIKILDPFPDHVVFKDLRITQQKI 1581

Qy        321 VALGGBFKLYLPQLIPHMLRVFMHDNSP-GRIVSIK-LLAAIQLFGANLDDYLH----  372
              |     |::|                   :    |:  || :|    | ||:|
Db       1582 KYSRGPFSLL--EEINHFLSVSVVDALPLTRLEGLKDLRRQLELHKDQMVDIMEASQDNP 1639

Qy        373 ---LLLPIVKLFDAPEAPL--PSRKAALETVDRLTESLDFTDYASRIIHPIVRTLDQSP  427
                 ::  ::   |  |  |           : |   |  |  |: |:  ||||    :|
Db       1640 QDGIMVKLVVNLLQLSKWAINHTGBKEVLEAVGSCLGEVGPIDFSTIAI----------- 1688

Qy        428 ELRSTAMDTLSSLVFQLQGKKYQ---IFIPMVNKVLVRHRINHQRYDVLICRIVKG----  479
                :|: | | ::    :|  :: :|    :  : | : |: |  | |:|||:|
Db       1689 QHSKDASYTKALKLFB-DKELQWTFIMLTYLNNTLVEDCVKVRSAAVTCLKNILATKTGH 1747

Qy        480 --YTLADEERDPLIYQHRMLRSGQGDALASGEVETGPMKKLHVSTINLQKAWGAARRVSK  537
                 : |     :  |  |: ::   |::        | :| :| :| |:|  |   |:|
Db       1748 SFWEIYKMTDPMLAYLQ---------PFRTSRKKFLEVPRFDKENPFEGL-----  1789

Qy        538 DD---WLE------WLRRLS-----------------LELLKDSSSPSLRSCWA----  565
              ||   |||      | :|:|                 |::|:|:   | |
Db       1790 DDINLWIPLSENHDIWKTLTCAFLDSGGTKCEILQLLKPMCEVKTDFCQYVLPYLHDI 1849

Qy        566 LAQAYNPMARDLFN----AAFVSC---------WSELNEDQQDELIRSIELALTSQDIA  611
              | | :||  || :|    ||  |           | |||  ::|:|  :|||::|||:|
Db       1850 LLQDTNESWRNLLSTHVQGFFTSCLRHPSQTSRSTTPANLDSESEHFFRCCLDKKSQ-- 1906

Qy        612 EVTQTLLNLAEFMEHSDKGPLPLRDDNGIVL----------LGERAAKCRAYAKALHY  659
              ::  :|| | :|     |  |           :          | |:|  : | |:|:
Db       1907 ---RIMLAVVDYMRRQK------RPSSGTIFNDAFWLDNLYLEVAKVAQSCAAHFTALLY 1957

Qy        660 KE--------LEFQKGPTPAILBSLISINNKLQQPEAAAGVLEVAMKHFGELE  704
              |:        :|  |||  :||:|  |:::|  |     || : | | ||  :
Db       1958 AEIYADKKSMDDQEKRSLAFEEGSQSTTISSL-SEKSKEETGISLQDLLLEIYRSIGEPD 2016

Qy        705 -IQATWYEKLHEWEDEALVAYDKKMDTNKDDPELMLGRMRCLEALG-------  748
               : |||:|  :| | :|:|  || || | |:| | :  | :|:||
Db       2017 SLYGCGGGKMLQPITRLRTVEHEAMWGKALVYD--LEFAIPSSTRQAGIIQALQNLGGLC 2074
```

```
Qy   1470 ANDPTSLRKNLSIQRYAVIPLSTNSGLIGWYPHCDTL-HALIRDYREKKILLNIEHRIM 1528
               :||:||  :||||| ||  ||        || |  |:  ||  | |   | ::
Db   2106 RKDQEANRRNLCIRTVVVIPLNEECGFIEWVNHTRPFREILLKSYRQKNIPISYQEHKVD 2165

Qy   1529 LRMAPPYDHITLMQKVFEHAVNNTAGDDLAKL-----LWL--KSPSSEVWFDRRTNYT 1581
               :  :         ||  |      |||            |||||||:|||||||
Db   2166 LDFA------LRSRNPGIFEKKI-------LPKFPPVFYEWFVESPFEPNNWVTSRQNYC 2213

Qy   1582 RSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGDCFRVAMTREKFPEKIPFRLTR 1641
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2214 RTLAVMSIVGYVLGLGDRHGENILFDEFTGEAIHVDFNCLFDKGLITFEK-PEKVPFRLTH 2272

Qy   1642 MLTNAMEVTGLDGNYRITCHTVMEVLRERHKDSVMAVLRAFVYDPLLNWRLMDTNTKGNKR 1701
               || | |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2273 NMVDAMGPTGYEGGFRKASEITMRLLRSNQDTIMSVLESFLHDPLVEWN------------ 2321

Qy   1702 SRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFIGDGLVKPEALNKKAIQII 1761
               ||                                    ||  :::
Db   2322 ----------------------------RKKSSSKYP----------------NNEANEVL 2338

Qy   1762 NRVRDKLTGRDFSHDDT--LDVPIQVELLIKQATSHENLCQCYIGWCPFW 1809
               ||||||     |||  ||  ||  | |:||||||||| |:| |::
Db   2339 DIIRKKFQG--FMPGETIPLSIEGQIQELIKSAVNPKNLVEMYIGWAAYF 2386

RESULT 9
ATM_HUMAN
ID   ATM_HUMAN    STANDARD;     PRT;   3056 AA.
AC   Q13315; O15429; Q12758; Q16551; Q93003; Q9NP02; Q9UCX7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Serine-protein kinase ATM (EC 2.7.1.37) (Ataxia telangiectasia
DE   mutated) (A-T, mutated).
GN   ATM.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96154672; PubMed=8589678;
RA   Savitsky K., Sfez S., Tagle D.A., Ziv Y., Sartiel A., Collins F.S.,
RA   Shiloh Y., Rotman G.;
RT   "The complete sequence of the coding region of the ATM gene reveals
RT   similarity to cell cycle regulators in different species.";
RL   Hum. Mol. Genet. 4:2025-2032(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97343327; PubMed=9199932;
RA   Platzer M., Rotman G., Bauer D., Uziel T., Savitsky K., Bar-Shira A.,
RA   Gilad S., Shiloh Y., Rosenthal A.;
RT   "Ataxia-telangiectasia locus: sequence analysis of 184 kb of human
RT   genomic DNA containing the entire ATM gene.";
RL   Genome Res. 7:592-605(1997).
RN   [3]
RP   SEQUENCE OF 1-24 FROM N.A.
RX   MEDLINE=97263790; PubMed=9108147;
RA   Savitsky K., Platzer M., Uziel T., Gilad S., Sartiel A., Rosenthal A.,
RA   Elroy-Stein O., Shiloh Y., Rotman G.;
RT   "Ataxia-telangiectasia: structural diversity of untranslated sequences
RT   suggests complex post-transcriptional regulation of ATM gene
RT   expression.";
RL   Nucleic Acids Res. 25:1678-1684(1997).
RN   [4]
RP   SEQUENCE OF 1-1369 FROM N.A., AND VARIANT AT 2546-SER--ILE-2548 DEL.
RX   MEDLINE=96381439; PubMed=8789452;
RA   Byrd P.J., McConville C.M., Cooper P., Parkhill J., Stankovic T.,
RA   McGuire G.M., Thick J.A., Taylor A.M.R.;
RT   "Mutations revealed by sequencing the 5' half of the gene for ataxia
RT   telangiectasia.";
RL   Hum. Mol. Genet. 5:145-149(1996).
RN   [5]
RP   SEQUENCE OF 1-2756 FROM N.A., AND VARIANT MCL LYS-750.
RA   Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
RA   Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA   Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 1349-3056 FROM N.A., AND VARIANT ASN-3003.
RX   MEDLINE=96105020; PubMed=8521392;
RA   Rasio D., Negrini M., Croce C.M.;
RT   "Genomic organization of the ATM locus involved in ataxia-
RT   telangiectasia.";
RL   Cancer Res. 55:6053-6057(1995).
RN   [7]
RP   SEQUENCE OF 1349-3056 FROM N.A., AND VARIANTS AT 2427-LEU-ARG-2428
RP   DEL; 2546-SER--ILE-2548 DEL AND SER-2860 DEL.
RC   TISSUE=Fibroblast;
RX   MEDLINE=9312868; PubMed=7792600;
RA   Savitsky K., Bar-Shira A., Gilad S., Rotman G., Ziv Y., Vanagaite L.,
RA   Tagle D.A., Smith S., Uziel T., Sfez S., Ashkenazi M., Pecker I.,
RA   Frydman M., Harnik R., Patanjali S.R., Simmons A., Clines G.A.,
RA   Sartiel A., Gatti R.A., Chessa L., Sanal O., Lavin M.F.,
RA   Jaspers N.G.J., Taylor A.M.R., Arlett C.F., Miki T., Weissman S.M.,
RA   Lovett M., Collins F.S., Shiloh Y.;
RT   "A single ataxia telangiectasia gene with a product similar to PI-3
RT   kinase.";
RL   Science 268:1749-1753(1995).
RN   [8]
RP   PARTIAL SEQUENCE FROM N.A., AND VARIANTS CYS-49; ARG-1054; PHE-1420;
RP   ILE-2079 AND ALA-2287.
RX   MEDLINE=96275738; PubMed=8665503;
RA   Vorechovsky I., Rasio D., Luo L., Monaco C., Hammarstroem L.,
RA   Webster A.D.B., Zaloudik J., Barbanti-Brodano G., James M.R.,
RA   Russo G., Croce C.M., Negrini M.;
RT   "The ATM gene and susceptibility to breast cancer: analysis of 38
RT   breast tumors reveals no evidence for mutation.";
RL   Cancer Res. 56:2726-2732(1996).
RN   [9]
RP   PHOSPHORYLATION.
RX   MEDLINE=97126018; PubMed=8969240;
RA   Chen G., Lee E.Y.-H.P.;
RT   "The product of the ATM gene is a 370-kDa nuclear phosphoprotein.";
RL   J. Biol. Chem. 271:33693-33697(1996).
RN   [10]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=97203148; PubMed=9050866;
RA   Brown K.D., Ziv Y., Sadanandan S.N., Chessa L., Collins F.S.,
RA   Shiloh Y., Tagle D.A.;
RT   "The ataxia-telangiectasia gene product, a constitutively expressed
RT   nuclear protein that is not up-regulated following genome damage.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:1840-1845(1997).
RN   [11]
RP   SUBCELLULAR LOCATION, AND VARIANTS AT 2546-SER--ILE-2548 DEL AND
RP   TYR-2824.
RX   MEDLINE=97294602; PubMed=9150358;
RA   Watters D., Khanna K.K., Beamish H., Birrell G., Spring K., Kedar P.,
RA   Gatei M., Stenzel D., Hobson K., Kozlov S., Zhang N., Farrell A.,
RA   Ramsay J., Gatti R.A., Lavin M.F.;
RT   "Cellular localisation of the ataxia-telangiectasia (ATM) gene product
RT   and discrimination between mutated and normal forms.";
RL   Oncogene 14:1911-1921(1997).
RN   [12]
RP   KINASE ACTIVITY.
RX   MEDLINE=97141775; PubMed=8988033;
RA   Jung M., Kondratyev A., Lee S.A., Dimtchev A., Dritschilo A.;
RT   "ATM gene product phosphorylates I kappa B-alpha.";
RL   Cancer Res. 57:24-27(1997).
RN   [13]
RP   C-ABL BINDING.
RX   MEDLINE=97311400; PubMed=9168117;
RA   Shafman T., Khanna K.K., Kedar P., Spring K., Kozlov S., Yen T.,
RA   Hobson K., Gatei M., Zhang N., Watters D., Egerton M., Shiloh Y.,
RA   Kharbanda S., Kufe D., Lavin M.F.;
RT   "Interaction between ATM protein and c-Abl in response to DNA
```

```
DR   EMBL; Y09076; CAA70297.1; -.
DR   EMBL; U76307; AAC49607.1; -.
DR   EMBL; AL049558; CAB40165.1; -.
DR   EMBL; X63544; CAA45106.1; -.
DR   PIR; T39911; T39911.
DR   GeneDB_SPombe; SPBC16.05; -.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008941; TPR-like.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; FALSE_NEG.
DR   PROSITE; PS00916; PI3_4_KINASE_2; FALSE_NEG.
DR   PROSITE; PS00290; PI3_4_KINASE_3; 1.
KW   Transferase; Kinase; DNA damage; DNA synthesis; DNA repair;
KW   Nuclear protein.
FT   DOMAIN    2018   2326    PI3K/PI4K.
FT   CONFLICT   198    198    T -> S (IN REF. 1).
FT   CONFLICT   526    526    L -> P (IN REF. 1).
FT   CONFLICT  1779   1781    SEC -> LRM (IN REF. 3).
SQ   SEQUENCE  2386 AA;  273526 MW;  DDEA343FFA3B72D3 CRC64;

Query Match            7.7%;    Score 725.5;  DB 1;  Length 2386;
Best Local Similarity  19.6%;   Pred. No. 5.2e-36;
Matches  410;  Conservative  327;  Mismatches  714;  Indels  639;  Gaps  84;

Qy    37 KALILKLKDPDPDENPGVINNVLATIGELAQVSGLEMRKWVDELFIIIMDMLQD--SSLL    94
         :|| |||:    :|    |||| ||| ||  ||  : ::|  :
Db   619 KALIFNTLKIIAVENTAILETVILSWISLSRVVEBE------ELHFVLLEVLSSVINSGi   672

Qy    95 AKRQVALWTLGQLVASTGYV-------VEPYRKYPTLLEVLLNFLKTEQN---------   137
Db   673 FYQGIGLSAL-QQIASTRHISVWQLLSPY--WPTVSGVAIVQGMGKKPNIASLFAQLMNIS   729

Qy   138 QG----------TRRBAIRV-LGLLGALD---------PYKHKVNIGMI--DQSRDA    172
Db   730 EGDFLIRTQAYTLPFLVLTKNKALIVRIAELSQSDVATLCLTNMHKILASLLTTDHPNLE   789

Qy   173 SAVSLSESKSSQDSSDYSTSEML----------VNMGNLPLDEFYPAVSMVALM--R    217
Db   790 ESVMLLSLATSDPEKVDLTSILRSDPISITVELLQLYQNDVPHEKIENALRKVAMIVSQ   849

Qy   218 IFRDQSLSHHHTMV-----VQAITFIFKSL-GLKCVQFLPQVMPTFLNV--IRVCDG    266
Db   850 VVNDEDLSNKELLYDFFNNHILGIIAEFSNILNDLKGKTSINEKIKTIVGIEKMLSLCGG   909

Qy   267 AIREPLFQQLGMLVS-FVKSHIR----------------------------PYM    291
Db   910 AVKLGLPQILSNLQSAFQNEHLRFVAIKAWFSLILLATKEPEYSSIAGLSLVILPLFPVL   969

Qy   292 D---EIVTLMREF---------------WVMNTSIQSPII-LLIEQIVVALGGE-FKLYLP   332
Db   970 EPQEAAELVIQIPDFISSDTHKCLQGLKWAIPTSLDSACFSLKAKEIFCSLQNEDFYSELQ  1029

Qy   333 QLIPHMLRVFMHDNSPGRIVSIKLLAA-------IQLFGANLDDYLHLLLPPI    378
Db  1030 SII---KCLTNENEPVCYLGLQKLELFFQAKVDELHDTLNLDISNEVLDQLLRCLLDCC   1085

Qy   379 VKLFDAP----------EAPLPSRKAA--------LETVDRLTESLDF-TDY-AS    413
Db  1086 VKYASTNMQISYLAAKNLGELGAIDPSRAKAQHIIKETVVLDNFENGEESLKFILDFMQS   1145

Qy   414 RIIHPIVRTLDQSPE--------LRSTAMDTLSSLVF------QLGKKYQ    449
Db  1146 QLIPAFLVTTDTKAQGFLAYALQEFLKLGGFKSAVINKKKGLTVVTEHWMSLPDLSKR-   1203

Qy   450 IFIPM------VNKVLVRHRINHQRYDVLICRIVKGYTLADEEEDPLIYQHRMLRSG    500
Db  1204 VLIPFLTSKYHLTPIPKIDIRYPI--YKENVTIHTWMQLFSLKLME-----YAHSQ----   1252

Qy   501 QGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDWLE------------WLRRL    547
Db  1253 ----------------------------NAEKIFGICSKVVKDQEVNIPCFLLPFLVLNVILTES   1289

Qy   548 SLE-------LLKDSSPSLRSCWALAQAYNPMARDLFNAA-FVSCWSEL---NED    592
Db  1290 BLEVNKVIEBFQLVINQPGPDGLNSVG--QQRYTSFVDVFFKIVDYLNKWLRMRKKRNWD   1347

Qy   593 QQDELIRSIELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRA    652
Db  1348 RRSAIARKENRYMSVEDATSRESSISKVBSFLSRFPSKTL-----GIVSL-----NCGF   1396

Qy   653 YAKALHYKELEFQKGPTP-AILES----LISINNKLQQPEAAAGV-LEYAMKHFGELEIQ    706
Db  1397 HARALFYWEQHIRNATAPYAALESDYRVLQEIYAGIDDPDEIRAVSLNF--HDYSFDQQ   1453

Qy   707 ATWYEKLHEWEDALVAYDKKWDTNKDDPELMLGRMCLEALGEWGQLHQQCCEKWTLVND    766
Db  1454 LLLHENSGTWDSALSCYEIIIQKDPENKKAKIGLLNSMLQSGHYESL-VLSLDSFIINDN   1512

Qy   767 ETQAKMARWAAAAAWGLGQWDSMEEYTCMIPRDTHD------GAFYRAVIALHQDLFSLA    820
Db  1513 HEYSKMLNLGIEASWRSLSIDSLKK--CLSKSNLBSFEAKLGSIFYQ--YLRKDSFA--   1565

Qy   821 QQCIDKARDLLDAELTAMAGESYSRAYGAMVSCH-MLSELEEVIQYKLVPERREIIRQIW    879
Db  1566 -ELTERLQPIYVDAATAIANTGAHSAY---DCYDILSKLHAINDFSRIAE----------   1611

Qy   880 WERLQGCQRIVEDWQKILMVRSL-------------------VVSPH----------EDMRTWLKY    916
Db  1612 -----TDGIVSDNLDIVLRRRLSQVAPYGKFKHQILSTHLVGYEKFENTKKTAEIYLEI   1665

Qy   917 ASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAVMKNMWKSARKIDAFQH    976
Db  1666 ARISRXNGQFQRAFNAII------KAMDLDKPLATIE------HAQWWWHQQQHRKAISE   1713

Qy   977 MQHFVQTMQQQAQHAIATEDQQHKQEL-HKLMARCFLKLGEW----QLNLQGINESTIP   1030
Db  1714 LNFSLNNNMPDLVDEHEERPXNRKETLGNPLRGKVFLKLTKWLGKAGQLGLKDL-ETYYH   1772

Qy  1031 KVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKLRHASGANITNATT   1090
Db  1773 KAVEIYS--ECENTHYYLGH------HRVLMYEEEQKLPVNEQSERFLSGELVTR----   1819

Qy  1091 AATTAATATTTASTEGSNSBSEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFR   1150
Db  1820 --------------------------------IINEFGR   1826

Qy  1151 SISLSRGNNLQDTLRVLTLWFDYG-----HWPDVNEALVEGV---KAIQIDT---   1194
Db  1827 SLYYGTNHIYESMPKLLTLWLDFGAEELRLSKDDGEKYFREHIISSRKKSLELMSNVCR   1886

Qy  1195 ----WLQVIPQLIARIDTDRPLVGRLIHQLLITDIGRYHQLLIYLP   1245
Db  1887 LSMKIPQYFFLVALSQMISRVCHPNNKVKILEHIIANVVASYPGETIWQLMATIKSTSQ   1946

Qy  1246 ARHNAANKILKNMCEHS---------NTLVQQAMMVSBELIRVAILWEHMWHEGLEEA   1294
Db  1947 KRSLRGKSIINVL--HSRKLSMSSKVDIKALSQSAILIITEKLINL-----------   1989

Qy  1295 SRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEPAQEWCRKYMKSGNV   1354
Db  1990 -------------------------CNTRINSKSV   1999

Qy  1355 KDLTQAWDLYYHVFR------RISKQLPQJTSLELQYVSPKLLMCRDLEAVPGTYDPNQP   1409
Db  2000 K----MSLKDH-FRLSFDDPVDLVIPAKSFLDI------TLPAKDANRASHYPPPKTQP   2047

Qy  1410 IIRIQSIAPSLQVITSKQRPRKLTLMGSNGHFFVPLLKGHBDLRQDERVMQLFGLVNTLL   1469
Db  2048 --TLLKFEDEVDIMNSLQKPRKVVRGTDGNLYPFLCKPKDDLRKDARLMFNNLICKIL   2105
```

```
Qy    774  RMAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLAIHQDLFSLAQQCIDKARDLLDA   833
                :||.||.|||||.::.|:  |::| || |:  |.:| |::||||:.|
Db   1462  PLAAGAAWGLGEWDMLEQYISVMKPKSPDKEFFDAILYLFKNDYDNASKHILNARDLLVT  1521

Qy    834  ELTAMAGESYSRAYGAMVSCHMLSELEVIQYKLVP---ERRBIIRQIWERLQGCCRIV     890
           |::|:  ||:|||:|.  |:::.|:||:|:|:|     ||::|:|::  ::.|||::|
Db   1522  BISALLINESYNRAYSVIVRTQIITEFEBIIKYKQ.PPNSEKKLLHYQNLWTKRLLGCCKNV 1581

Qy    891  EDWQKILMVRSLVVSPHEDMRTWLKYASLGCKSGRLALAHKTLVILL--GVDPSRQLDHP   948
           :.||::|.||||:|  |:|:: |:|:|:||.|||:::|:||:|::||   |.|
Db   1582  DLWQRVLRYRSLVIVKPKQDLQIWIKFANLCRKSGRMRLANKALNMLLEGGNDPS------  1635

Qy    949  LPTVH---PQVTYAYMKQNWKSARKIDAFQRMQHFV------QTMQQQAQHAIAT       994
           ||.. |   |:|  :|:::|:: |:|:.||.::::|        ::.|:|::.:::
Db   1636  LPNTFKAPPPVYAQLKYIWATGAYKEALNHLIGFTSRLARHDIGLGDPNNMIAQSVKLSSA 1695

Qy    995  EDQQHKQELHKLMARCPLKLLGEWQLNLQGINESTIP-KVLQYYSAATEDHRSWYKAWHAW  1053
           | ::::|:|:||:|||.||  ||:::  |||  |: | ||| .:: ||.|::||||||:|
Db   1696  STAPYVEBYTKLLARCFLKGGEWRIATQPNWRNTNPDAILGSYLLATHFDKNWYKAWHNW   1755

Qy   1054  AVMNFEAVLHYKHQNQQARDEKKKLRHASGANITNATTAATAATATTTASTEGSNESSEA  1113
           |:.|||.|:  :.   ::.|.   |:.|.|||:  |:::|:||::|
Db   1756  ALANFEVI------SMVQETKLN--GGKNDDDDTAVN-----NDNVRIDGSILGSGS     1801

Qy   1114  ESTENSPPPSPLQKKVTEDISKTILMVTPAVQGFFRSISLSRGNNLQDTLRVLTLWFDY   1173
           ::.||  |: |:|:|:|| |||::|.|   ||:|:  |::|:.:||||||:.|:|||:::
Db   1802  LTINGNRYPLELIQR----------HVVPAIKGFFHSISLLETSCLQDTLRLLTLLFNF   1850

Qy   1174  GHWPDVNRALVEGVKAIQIDTWLQVIFQLIARIDTPRPLVGRLIHQLLTDIGRYHPQALI  1233
           |.:::|::|::||.:::|||:||:|||||:||| || :.:::|:.||:|:||:|||||::
Db   1851  GGIKEVSQAMYEGFNLMKIENWLEVILFQLISRIHQPDPTVSNSLLSLLSDLGKAHPQALV 1910

Qy   1234  YPLTVAGKSTTTARENAANKILKNMCEHSNTLVQQAMMVSEBLIRVALIWHEMWHBGLEE  1293
           |||||| |:::.|..:|:.||:|::||:|| || ::|::|:.|||||::||:.|:.||:
Db   1911  YPLTVAIKSESVSRQKAALSITEKIRIHSPVLVNQAELVSHELIRVAVLWHELWYEGLED  1970

Qy   1294  ASRLYFGERNVKGMFEVLBPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGN  1353
           |||::|.|.::|:|:.|||||| .||:.|||||::||:::.|||||:||||  |:|:||:
Db   1971  ASRQFFVEHNIEKMFSTLRPLHKGLGNEPQTLSEVSFQKSFGRDLNDAYEWLNNYKKSKD  2030

Qy   1354  VKDLTQAWDLYYHVFRRISKQLFQLTSLELQYVSPKLLMCRDLELAVPGTYDNQPLIRI   1413
           ::::.||||:||:|||:|.:|:.||.:|:||:||| ||.::||||||||||  . |:||
Db   2031  INNLNQAWDIYYNVFRKIIRQIPQLQTLDLQHVSPQLLATHDLELAVPGTYFPGKPTIRI  2090

Qy   1414  QSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDP  1473
           ::.. |.||:|||||||:|::.||.|::.|::|||||:|||| ||||||||||||:.||
Db   2091  AKFPPLFSVISSKQRPRKFSIKGSDGKDYKYVLKGHEDIRQDSLVMQLFGLVNTLLKNDS  2150

Qy   1474  TSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYREKKKILLNIEHRIMLRMAP  1533
           :.::::| |||| |||||:.||::||::.||:|.|||::.:||:|.||:||::|| |||
Db   2151  ECFKRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREBHRDAKKIFLNIEHWVMLQMAP 2210

Qy   1534  DYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYI  1593
           ||::|||:||:|||.:|::||.|.||.|::|||| |||.|::||||||||||||||:|||
Db   2211  DYENLTLLQKIEVFTYALDNTKGQDLYKIIWLKSRSSETWLERRTTYTRSLAVMSWTGYI  2270

Qy   1594  LGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKPEKIPFRLTRMLTNAMEVTGLD  1653
           |||||||||||||||::||::||||||||:|:::||:|||:|||||||||:||||:|:
Db   2271  LGLGDRHPSNLMLDRITGKVIHIDFGDCFEAAILREKYPEKVPFRLTRMLTYAMEVSGIE 2330

Qy   1654  GNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKGRSRTRTDSYSAGQ 1713
           |.:|||| .|:.||::.|:|:||:||||. |||::|.   :.:|.
Db   2331  GSFRITCENVNRVLRDNKESLMAILEAFALDPLIHWGFDLPPQKLTBQT--------GI   2381

Qy   1714  SVEILDGVELGEPAHKKTGTTVPESIIHSFIGDGLVKPEALNKKAIQINRVRDKLTGRDF  1773
            ::|:  :|.:|    ::..:||.  ::..||:.| :|:.||  |:::: |:||||||:|
Db   2382  PLPLINPSEL----LRKGAITVEEAANM--EAEQQNETKNARAMLVLRRIITDKLTGNDI  2434

Qy   1774  SHDDTLDVPTCVELLIKQATSHENLCQCYIGWCPFW                         1809
           .:::.||||  |:||||||||.   |:.|||||||||
Db   2435  KRFNELDVPEQVDKLIQQATSIRRLCQHYIGWCPFW                         2470
```

RESULT 8

RAD3_SCHPO

ID   RAD3_SCHPO   STANDARD;   PRT;   2386 AA.
AC   Q02059; Q92391; Q9UUM1;
DT   01-JUL-1993 (Rel. 26, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   DNA repair protein rad3.
GN   RAD3 OR SPBC216.05.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=972;
RX   MEDLINE=97133293; PubMed=8978690;
RA   Bentley N.J., Holtzman D.A., Flaggs G., Keegan K.S., DeMaggio A.,
RA   Ford J.C., Hoekstra M., Carr A.M.;
RT   "The Schizosaccharomyces pombe rad3 checkpoint gene.";
RL   EMBO J. 15:6641-6651(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=2184401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
RN   [3]
RP   SEQUENCE OF 711-1781 FROM N.A.
RX   MEDLINE=93012978; PubMed=1398093;
RA   Seaton B.L., Yucel J., Sunnerhagen P., Subramani S.;
RT   "Isolation and characterization of the Schizosaccharomyces pombe rad3
RT   gene, involved in the DNA damage and DNA synthesis checkpoints.";
RL   Gene 119:83-89(1992).
CC   -!- FUNCTION: Involved in G2 arrest following DNA damage where it
CC   phosphorylates chk1. It is also involved in the dependence of
CC   mitosis on the completion of DNA replication.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.

```
RA  Huang M.-E., Manus V., Chuat J.-C., Galibert F.;
RT  "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT  reading frames and a gene cluster with a counterpart on chromosome
RT  XI.";
RL  Yeast 12:869-875(1996).
CC  -!- FUNCTION: Phosphatidylinositol 3-kinase homolog required for G1
CC      progression. Target of the antibiotic rapamycin.
CC  -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC      1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC  -!- MISCELLANEOUS: It may act on another substrate or phosphorylate a
CC      different position in the phosphatidylinositol ring.
CC  -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC  -!- SIMILARITY: Contains 12 HEAT repeats.
CC  -!- CAUTION: It is uncertain whether Met-1 is the initiator.
CC  --------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------------
DR  EMBL; L19540; AAB66881.1; -.
DR  EMBL; X74857; CAA52849.1; -.
DR  EMBL; Z49566; CAA89594.1; -.
DR  EMBL; L47993; AAB39292.1; -.
DR  PIR; S57085; S57085.
DR  HSSP; P42345; 1FAP.
DR  GermOnline; 141899; -.
DR  SGD; S0003827; TOR1.
DR  GO; GO:0010008; C:endosome membrane; IDA.
DR  GO; GO:0000139; C:Golgi membrane; IDA.
DR  GO; GO:0005886; C:plasma membrane; IDA.
DR  GO; GO:0005774; C:vacuolar membrane; IDA.
DR  GO; GO:0005515; F:protein binding; IPI.
DR  GO; GO:0000080; P:G1 phase of mitotic cell cycle; IMP.
DR  GO; GO:0007126; P:meiosis; IMP.
DR  GO; GO:0000074; P:regulation of cell cycle; IMP.
DR  GO; GO:0007046; P:ribosome biogenesis; IMP.
DR  InterPro; IPR008938; ARM.
DR  InterPro; IPR003151; FAT.
DR  InterPro; IPR003152; FATC.
DR  InterPro; IPR003357; HEAT.
DR  InterPro; IPR000403; PI3_PI4_kinase.
DR  Pfam; PF02259; FAT; 1.
DR  Pfam; PF02260; FATC; 1.
DR  Pfam; PF00454; PI3_PI4_kinase; 1.
DR  SMART; SM00146; PI3Kc; 1.
DR  PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
DR  PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR  PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR  PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW  Transferase; Kinase; Cell cycle; Repeat.
FT  REPEAT        114    151       HEAT 1.
FT  REPEAT        249    286       HEAT 2.
FT  DOMAIN        441    447       ARG/LYS-RICH (BASIC).
FT  REPEAT        627    663       HEAT 3.
FT  REPEAT        664    701       HEAT 4.
FT  REPEAT        701    784       HEAT 5.
FT  REPEAT        788    826       HEAT 6.
FT  REPEAT        832    870       HEAT 7.
FT  REPEAT        870    908       HEAT 8.
FT  REPEAT        908    946       HEAT 9.
FT  REPEAT        950    987       HEAT 10.
FT  REPEAT       1069   1107       HEAT 11.
FT  REPEAT       1109   1147       HEAT 12.
FT  REPEAT       1870   1907       FAT.
FT  DOMAIN       2125   2470       PI3K/PI4K.
FT  MUTAGEN      1972   1972       S->R,N: LOSS OF RAPAMYCIN SENSITIVITY.
FT  CONFLICT       58     58       D -> G (IN REF. 1).
FT  CONFLICT      115    115       V -> I (IN REF. 1).
FT  CONFLICT      133    133       S -> N (IN REF. 1).
FT  CONFLICT      231    231       A -> R (IN REF. 2).
FT  CONFLICT      396    396       N -> K (IN REF. 1 AND 2).
FT  CONFLICT      547    547       N -> S (IN REF. 1 AND 2).
FT  CONFLICT      675    675       T -> I (IN REF. 2).
FT  CONFLICT     1292   1292       G -> E (IN REF. 2).
FT  CONFLICT     1436   1436       G -> A (IN REF. 2).
FT  CONFLICT     1468   1468       A -> R (IN REF. 1 AND 2).
FT  CONFLICT     1469   1471       WGL -> GGS (IN REF. 2).
FT  CONFLICT     1478   1479       EQ -> DE (IN REF. 2).
FT  CONFLICT     1590   1590       V -> I (IN REF. 2).
FT  CONFLICT     1632   1642       NDPSLPNTFKA -> TILVVQIRSKP (IN REF. 2).
FT  CONFLICT     1640   1640       F -> Y (IN REF. 1).
FT  CONFLICT     1844   1844       L -> S (IN REF. 1).
FT  CONFLICT     1972   1972       S -> R (IN REF. 1).
FT  CONFLICT     2202   2202       H -> Q (IN REF. 1).
FT  CONFLICT     2414   2414       K -> R (IN REF. 1 AND 2).
SQ  SEQUENCE   2470 AA;  281136 MW;  ACB1781B9963BB1E CRC64;

Query Match             42.1%;  Score 3962;  DB 1;  Length 2470;
Best Local Similarity   44.4%;  Pred. No. 3e-235;
Matches  815;  Conservative 360;  Mismatches 551;  Indels 110;  Gaps 26;

Qy          1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALLLKLKDPDPDPNPGVINNVLA 60
              |||||:|:|||:||||: |||:|: | ||||| || |:|:|| :| :||        |
Db        718 LKFSTSSREKEETASLLCTLIRSSKDVAKPYIEPLLNVLLPKFQ---DTSSTVASTAIR 773

Qy         61 TIGELAQVSGLEMRKKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
              |||||  : | ||| ||| ::|| |:||||  |  |::|:|| |||||:| :|||:||
Db        774 TIGELSVVGGEDMKIYLKDLFPLIIKTFQDQNSFKREALKALGQLAASSGYVIDPLLD 833

Qy        121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
              ||::||:|||:|| |  |  |    :|| || :|:||:|:   ||    :  |   ::
Db        834 YPELLGILVNILKTENSQNIRRQTVTLIGILGAIDPYRQK----------EREVTSTTDI 883

Qy        181 KSSQDSDYSTSEMLVNMGNLPL-DEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIF 239
              | |  || : |:|  |||| || :|||||| | :| :|:||||:|||:|:| || :|
Db        884 STEQNAPPIDIA--LLMQGMSPSNDEYYTVVHCLLKILKDPSLSSYHTAVIQAIMHIF 941

Qy        240 KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR 299
              ||| ||:| || :|| | |::: |: || |||| ||||:||:|:|:|||||||:|:|:|:
Db        942 QTLGLKCVSFLDQIIPTILDVMRTCSQSLLEYFQQLCSLIIIVRQHIRPHVDSIFQAIK 1001

Qy        300 BFWVMNTSIQSTIILLIEQIVALGGBEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAA 359
              :  ::|: |:|:|||||: ::|  :|||:||| |:|| |:  ||:| |||||:|||||
Db       1002 DFSSV-AKLQITLVSVIEAISKALEGEFKRLVPLILTLFLVILENDKSSDKVLSRRVLRL 1060

Qy        360 IQLFGANLDDYLHLLLPPIVKLFDPAPEAPLPSRKAALETVDRLTESLDFTDYARIIHPI 419
              :||| |||:| |||:||||:||||  | | | || :||:|    |:  |:|||||||::|
Db       1061 LESFGPNLEGYSHIITPKIVQMAEFTSGNL-QRSAIITIGKLAKDVDLFEMSSRIVHSL 1118

Qy        420 VRTLDQ--SPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIV 477
              |:|| :  |:|:|::|:| |:|::|||| || ||:|::|:|:|:|::|:||||:|:||
Db       1119 LRVLSSTTSDELSKVIMNTLSLLIQMGTSFAIFIPVINEVLMKKHIQHTIYDDLTNRIL 1178

Qy        478 KGYTLADEEBDPLIYQHRMLRSGQGDALASGPVETGP--MKKLHVSTINLQKAWGAARRV 535
              ||: |||||:|| |  |   :     ||  || |:|   |:|||:||:|||:| | | |:
Db       1179 NNDVLPT--------KILEANTDYKPAEQMEARADAGVAKLPINQSVLKSAWNSSQQR 1228

Qy        536 SKDDWLEWLRRLSLELLKDSSSPSLRSCWAIAQAYNPMARDLFNAAFVSCWKSELNEDQQD 595
              |||| ||||||| |:|||||||||||||| ||:|||| | ||| || || ||| || :|
Db       1229 TKEDWQEWSKRLSIQLLKESPSHAIRACSNLASMYYPLAKELPNTAFACVWTELYSQVQE 1288

Qy        596 ELIRSIELALTSQ-DIAEVTQTLLNLAEFMEHSDKG-PLPLRDDNGIVLLGERAAKCRAY 653
              :|:|| |:| | | |: ||||:|||||| ||:| ||  | | || :||||||:|  ||||
Db       1289 DLIGSLCIALSSPLNPPEIHQTLNLVEFMEHDDKALPIPTQS------LGEYAERCHAY 1342

Qy        654 AKALHYKELBFPQKGPTPAILESLLSINNKLQQPEAAAGVLEYAMKHPGELEIQATWYEKL 713
              |||||||||: |:|||| ||||||||:|||||:| :||| || ::|  :|::||  ||||
Db       1343 AKALHYKEIKFIKEPENSTIESLISINNQLNQTDAAIGLKHAQQHH-SLQLKETWFEKL 1401

Qy        714 HEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMA 773
              : ||||: |  | |: |  :| | |::   ||| ||||||||||| | |:|:||||||||
Db       1402 ERWEDALHAYNEREKAGDTSVSVTLGKMRSLHALGEWEQLSQLAARKWKVSKLQTKKLLA 1461
```

```
Qy  300  EFWMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMEHNSPGRIVSIKLLAA  359
         |:: :||||| :|| ||   :: :||   : : ::||  : :||:
Db 1011  EFFPI-IKLJQTFIISVIESISKALEGBEFKRFVPETLTFFLDILENDQSNKRIVPIRLLKS 1069

Qy  360  IQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPI  419
         : |||:|||:|  |:||||:|: :: ::  | ||  ::| :||||
Db 1070  LVTFGPNLEDYSHLIMPIVVRMTEYSAGSL--KKISIITLGRLAKNINLSEMSSRIVQAL  1127

Qy  420  VRTLDQSP-ELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVK  478
         || |   | ||  :|::||||||:||| ||||  :::|  ||:||:||
Db 1128  VRILNNGDRELTKAIMNTLSLLLQLGTDFVVFVPVINKALLRNRIQHSVYDQLVNKLLN  1187

Qy  479  GYTLA------DEEBDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAA  532
         ::                 :  ::|                :| : ||:::
Db 1188  NECLPTNIIFDKENEVP-------ERKNYEDEM------QVTKLPVNQNILKNAWYCS  1232

Qy  533  RRVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNED  592
         :: :: ||:||:|||:::|||: ::| :|| | : :||: ||::||:| :||| ||:
Db 1233  QQKKXEDWQEWIRRLSIQLLKESPSACLRSCSSIVSVYPLARELFNASPSSCWVELQTS  1292

Qy  593  QQDELIRSIELAL-TSQDIABVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCR  651
         :|:||| :|  || :| |  : || :|||:|||||::| |||   |  :|||:|| ||
Db 1293  YQEDLIQALCKALSSSENPEIYQMLNLVBFMEHDDK-PLPI----PIHTLGKYAQKCH  1347

Qy  652  AYAKALHYKELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYE  711
         |:||||||||:||  ::  ||| |||||::|:| :| | ::|:||||| :| |||:||||
Db 1348  AFAKALHYKEVEFLEEPKNSTIEALISINNQLHQTDSAIGIIKHAQQH-NELQLKETWYE  1406

Qy  712  KLHSWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAK  771
         || :|||||.| :|        | | :|:|::|||||||::  ||: |  : |:|:
Db 1407  KLQRWEDALAAINEKBAAGEBDSVEVVMGKLRSLYALGEWEELSKLASEKWGTAKPEVKKA  1466

Qy  772  MARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLL  831
         ||  ||||:|||:|||: |:|| |:    |:|:|| |:|| ||:   |  | |:|||||
Db 1467  MAPL-AAAAWGLEQWDEIAQYTSVMKSQSPDKEFYDAILCLHRNNFKKAEVHIFNARDLL  1525

Qy  832  DAELTAMAGESYSRAYGAMVSCHMLSELEEVIQYKLVP--ERRBIIRQIWERLGGCQR    888
         ||| : |: |||:||:. :| : |:|:||||:::|| | ::  :|: : |  |||:|:
Db 1526  VTELSALVNESYNRAYNVVVRAQIIAELEBIIKYKKLPQNSDKRLTMRETWNTRLLGCCQK  1585

Qy  889  IVEDWQKILMVRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHP  948
         ::|:|| ||:||:|:: : ::|  ||| |:||::||:|||:|:|: || :: : |||||
Db 1586  NIDVWQRILRVASLVIVKPKEDAQVRIKFANLCRKSGRMALAKKVLNTLL--EETDDPDHP  1643

Qy  949  -LPTVHPQVTYAYMKNMWKSARKIDAFQEMQHF----------------VQTMQQQAQHA  991
          :|  :|  | |:::: ::|||| ||  |                    ||::| |  |
Db 1644  NTAKASPPVVYAQLYLWATGLQDEALKQLINFTSRMAHDLGLDPNNMIAQSVPQQSKRV   1703

Qy  992  IATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPK-------VLQYYSAATEHD 1043
          : ||:| :|:| |:|||||| |||::  |:   :|:|         |:  | :|  |
Db 1704  -----PRIVEDYTKLLARCFLKQGEWRVCLQ----PKWRLSNPDSIIGSYILATHFD   1751

Qy 1044  RSWTKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTATATATTAS     1103
         .:|.|||| ||: ||:||:|||:                          ::  ::
Db 1752  NTWYKAWHNWALANFVI---------------------------SMLTSVSKKK      1779

Qy 1104  TEGGNSBEEASTE--NSPTPSPLQKKVTEDLSKTLLMYTTVPAVQGFFRSISLSRGNNL  1160
         .:|:||  |||.|  ||.|.| |::|:.|:|| |:| |::: || |||:|||||.::|
Db 1780  QEGSDASSVTDINEFDNGMIGVNTFDAKEVHYSSNLIHRHVIPAIKGFFHSISLSESSSL  1839

Qy 1161  QDTILRVLLIWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRPLVGRLIHQL 1220
         ||::|::|| :|| |: | :.|| ::|::|| |::||:|: |||:||:||  |::|||: |
Db 1840  QDALRLILTMFTWFFGGIPEATQAMEEGFNLIIQIGTWLEVLEVLPQLISRIHQPNQIVSRSLLSL 1889

Qy 1221  LTDIGRYHPQALIYLPLTVASKGTTTARHNAANKILKRMCEHSNTLVQQAMMVSEELLRVA 1280
         |:|:|: |||||:| ||.|: ||::: ::::|::|| :|| ||:.||::|  ||||:||:|
Db 1890  LSDLGKAHPQALVYPLMVAIKSSSLSRQKAALSIIEKMRIHSPVLVDQAELVSHELIRMA  1959

Qy 1281  ILWHEWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLME  1340
         :|||| :||| :||::.||: |  :||  ||||:.|::|||:|| ::||:: :|||||
Db 1960  VLWHEQWYEGLDDASQRFGEHNTEKMFAALEPLYEMLKRGPETLREISFQNSFGRDLND  2019

Qy 1341  AQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAV  1400
         | | .::::||:|||:| :|||||||:|| :: ||||: ||||:|||||| .:||:||||
Db 2020  AYEKLMNVKKSKDVSNLNQAWDLYYNVFRKIGKLPQLQTLELQHVSPKLLSAHDLELAV   2079

Qy 1401  PGT-YDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVM  1459
         ||    ...  ||| |   : :: ||||.|||  : ||.| : ::|||||:||||: ||
Db 2080  PGTRASGGKPIVKISKFPEVFSVISSKQPRKFCIKGSDGKDYKVLKGHEDIRQDSLVM    2139

Qy 1460  QLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYREKKKI  1519
         ||||||||||:||    |:.| |||| |||| :.||:|||  :|| |:||: :||:|||
Db 2140  QLFGLVNTLLQNDABCFRRHLDIQQYPAIPLSPKSGLLGWVPNSDTFHVLIREHREAKKI  2199

Qy 1520  LLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTN  1579
          ||||: :||:|||||| |||:||||| :|::||| |:||.|:||:|| || ||:.||:
Db 2200  PLNIEEHWVMLQMAPDYDNLTLLQKVEVFTYALNNTEGQDLYKVLWIKSRSSETWLERRTT  2259

Qy 1580  YTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPFRL  1639
         |||||||||| |||||||||||||||||| :|| :|||||||||.|: |||||||:|||
Db 2260  YTRSLAVMSMTGYILGLGDRHPSNLMLDRITGKVIHIDFGDCFEAAILREKFPEKVPFRL  2319

Qy 1640  TRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGN  1699
         ||||| |||| |::|  ||||  ||:||:::| |:||:||||  |||:||  :|   |
Db 2320  TRMLTYAMEVSGIEGSFRITCENVMKVLRDNKGSLMAILEAFAFDPLINWG-FDLPTK--  2376

Qy 1700  KRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVP-ESIHSFIGDGLVKPEAL-----  1753
                                         |: :||
Db 2377  ------------------------------KIEEETGIQLPVMNANELLSNGAITEEEVQRVEN 2410

Qy 1754  ------NKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQCYIGWC  1806
               ::||| ::||. :|||||| |  ||:|||:||:|:|||||||:||||| |||||
Db 2411  EHKNAIRNARAMLVLKRITDKLTGNDIRRPNDLDVPEQVDKLIQQATSVENLCQHYIGWC  2470

Qy 1807  PFW  1809
         |||
Db 2471  PFW  2473
```

RESULT 7
TOR1_YEAST
ID   TOR1_YEAST   STANDARD;   PRT;   2470 AA.
AC   P35169;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Phosphatidylinositol 3-kinase TOR1 (EC 2.7.1.137) (PI3-kinase)
DE   (PtdIns-3-kinase) (PI3K).
GN   TOR1 OR DRR1 OR YJR066W OR J1803.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1].
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94019276; PubMed=8413204;
RA   Cafferkey R., Young P.R., McLaughlin M.M., Bergsma D.J., Koltin Y.,
RA   Sathe G.M., Faucette L., Eng W.-K., Johnson R.K., Livi G.P.;
RT   "Dominant missense mutations in a novel yeast protein related to
RT   mammalian phosphatidylinositol 3-kinase and VPS34 abrogate rapamycin
RT   cytotoxicity.";
RL   Mol. Cell. Biol. 13:6012-6023(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JK-3D;
RX   MEDLINE=94243030; PubMed=8186460;
RA   Helliwell S.B., Wagner P., Kunz J., Deuter-Reinhard M., Henriquez R.,
RA   Hall M.N.;
RT   "TOR1 and TOR2 are structurally and functionally similar but not
RT   identical phosphatidylinositol kinase homologues in yeast.";
RL   Mol. Biol. Cell 5:105-118(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=96437976; PubMed=8840504;

```
Qy 1354 VKDLTQAWDLYYHVFRRISKQLPQLTSLELQYVSPKLLMCRDLELAVPGTYDPNQPIIRI 1413
         :|:||:|||:::||:|||||  ||:||||||||||||   ||||||||||| :|:|||
Db 1893 PTELNQSWDLYQVFRRIQKQLPRIKHLELQYVSPKLLDACDLELAVPGTYGHNKPVIRI 1952

Qy 1414 QSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFGLVNTLLANDP 1473
         : :|: |||:|||||:| |:.: :|: | |||||||||||||||||||| ||||:|
Db 1953 SHFHHTFEVISSKQRPRRLTIHGSDGKDYQYVLKGHEDLRQDERVMQLFGLCNTLLTTDS 2012

Qy 1474 TSLRKNLSIQRYAVIPLSTNSGLIGWVPHCDTLHALIRDYREKKKILLNIEHRIMLRMAP 1533
         :||:| ||:||:|||||:.|||:||||| |||| ||:|::|| |:|||||:||:||:||
Db 2013 ETPKRRLNIERYTVIPLSPNSGLLGWVPHSDTLHFLIKEFRSKRNILNLEHRMMLQMAP 2072

Qy 1534 DYDHLTLMQKVEVFEHAVNNTAGDDLAKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYI 1593
         | ..|||:|| |||:|| |:|.| ||: |||||:.||.|:||||:||:|||||||||||
Db 2073 DCDSLTLLQKLEVFEYVMANTDGYDLYHVLWLKSRSSEAWLDRRTSYTQSLAVMSMVGYI 2132

Qy 1594 LGLGDRHPSNIMLDRLSGKILHIDFDGCFEVAMTREKFPEKIPFRLTRMLTNAMRVTGLD 1653
         |||||||||| |:||.|||| |||| | |||||:|||||||||||||||| |||:|:|:
Db 2133 LGLGDRHPSNLMMDRYSGKIIHIDFPGDCFEVAMHREKFPEKIPFRLTRMLINAMEVSGIQ 2192

Qy 1654 GNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQ 1713
         |.|:||| .||.||| . ||:||||||||||||:|||| ||.       :||:||||
Db 2193 GTYKITCELVMRVLRSNTESLMAVLEAFVYDPLINWRLMLTYKSSFGASTLRPTSS------ 2247

Qy 1714 SVEILDGVELGEPAHKKTGTTVPESIHSFSIGDGLVKPEALNKKAIQINRVRDKLTGRDF 1773
         |||       |||.  ||:||:|:||:|  ||:|||| ||||:|::||||||||.||:||
Db 2248 SVE-------EKGRSYTHRARHADYAALSETNGVNAEGLNERSIQVLKRVSNKLTGKDF 2299

Qy 1774 SHDDTLDVPTQVELLIKQATSHENLCQCVIGWCPFW 1809
         .::|.| ||:|| |||||||:.|| |:|| ||||:|
Db 2300 DLKEQLPVKAQVEKLIQQATAPENJLCRCVGWCSFW 2335
```

RESULT 6
TOR2_YEAST
ID   TOR2_YEAST   STANDARD;   PRT;   2473 AA.
AC   P32600;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Phosphatidylinositol 3-kinase TOR2 (EC 2.7.1.137) (PI3-kinase)
DE   (PtdIns-3-kinase) (PI3K).
GN   TOR2 OR DRR2 OR YKL203C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JK9-3D;
RX   MEDLINE=93258821; PubMed=8387896;
RA   Kunz J., Henriquez R., Schneider U., Deuter-Reinhard M., Movva N.,
RA   Hall M.N.;
RT   "Target of rapamycin in yeast, TOR2, is an essential
RT   phosphatidylinositol kinase homolog required for G1 progression.";
RL   Cell 73:585-596(1993).
RN   [2]
RP   SEQUENCE OF 1-2360 FROM N.A.
RA   Maia e Silva A., Bossier P., Vilela C., Fernandes L., Soares H.,
RA   Guerreiro P., Rodrigues-Pousada C.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1902-2473 FROM N.A.
RA   Pohl T.M., Pohl F.M.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Phosphatidylinositol 3-kinase homolog required for G1
CC   progression. Target of the antibiotic rapamycin.
CC   -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC   1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   -!- SIMILARITY: Contains 12 HEAT repeats.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; X71416; CAA50548.1; -.
DR   EMBL; Z28203; CAA82048.1; -.
DR   PIR; S38040; S38040.
DR   HSSP; P42345; 1AUE.
DR   GermOnline; 139957; -.
DR   SGD; S0001686; TOR2.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0005774; C:vacuolar membrane; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0007010; P:cytoskeleton organization and biogenesis; IMP.
DR   GO; GO:0007266; P:Rho protein signal transduction; IMP.
DR   GO; GO:0007046; P:ribosome biogenesis; IMP.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000357; HEAT.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc_1.
DR   PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Transferase; Kinase; Cell cycle; Repeat.
FT   REPEAT   588   626   HEAT 1.
FT   REPEAT   636   674   HEAT 2.
FT   REPEAT   676   710   HEAT 3.
FT   REPEAT   756   793   HEAT 4.
FT   REPEAT   797   835   HEAT 5.
FT   REPEAT   841   879   HEAT 6.
FT   REPEAT   917   955   HEAT 7.
FT   REPEAT   1039   1076   HEAT 8.
FT   REPEAT   1079   1116   HEAT 9.
FT   REPEAT   1118   1155   HEAT 10.
FT   REPEAT   1292   1331   HEAT 11.
FT   REPEAT   1872   1909   HEAT 12.
FT   DOMAIN   2128   2473   PI3K/PI4K.
FT   CONFLICT   1472   1472   A -> AG (IN REF. 1).
SQ   SEQUENCE   2473 AA; 281508 MW; 42C73D65CDD4DB5F CRC64;

Query Match   43.6%; Score 4105; DB 1; Length 2473;
Best Local Similarity   45.5%; Pred. No. 4.7e-244;
Matches 847; Conservative 333; Mismatches 513; Indels 170; Gaps 28;

```
Qy   1 LEHSGIGRIKEQSARMLGHIVSNAPRLIRPYMEPILKALILKLKDPPDPNPGVINNVLA 60
         |.:| |  :|::.|: ||::  ||| :|:||::||||: :|   .|  :.  :.: :
Db 727 LKFSNMPKKKEBSATLLCTLINSSDEVAKPYIDPILDVILPKCQ---DASSAVASTALK 782

Qy  61 TIGELAQYSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK 120
         .:|| :|||||:|:|.||::|::||||:||| ||.|||||||||||||:|||||||||:|
Db 783 VLGELSVVGGKEMTRYLKELMELMLMDSLSIHHTAAIQAIMHIF--------------- 842

Qy 121 YPTLLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYHKVNIGMIDQQSRDASGAVSLSES 180
         || |:::|:|:||||:|  :|||:|:|||:|:|||||:||||| :|            :
Db 843 YPBELGIILNILKTENNPHIRRGTVRLIGIIGALDPYKHR------BIEVTSNS 890

Qy 181 KSSQDSSDYSYSTSEMLVNMGNLPL-DEFYPAVSMVALMRFRDQSLSHHTMVVQAITFIF 239
         |||  :. :.||:||:::.|.|:||  :.||:| ::. : .: :|||:|:||:|||::|||
Db 891 KSSVBQNAPSIDIALLMQGVSPSNDBYYPTVIHNLMKILNDPSLSIHHTAAIQAIMHIF 950

Qy 240 KSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR 299
         ::|||:||.||.|::|.::|:.|.| ::|.|||:||||:|:|::|.|||| :::|: :|:
Db 951 QNLGLRCVSFLDQIIPGIILVWRSCPPSQLDFYPQQLGSLIISIVQQHIRPHVEKIYGVIR 1010
```

```
CC       oxidative stress.
CC    -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC    -!- SIMILARITY: Contains 16 HEAT repeats.
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
DR    EMBL; Z97992; CAB10805.1; -.
DR    PIR; T40186; T40186.
DR    HSSP; P42345; 1FAP.
DR    GeneDB_SPombe; SPBC30D10.10c; -.
DR    InterPro; IPR008918; ARM.
DR    InterPro; IPR003151; FAT.
DR    InterPro; IPR003152; FATC.
DR    InterPro; IPR000357; HEAT.
DR    InterPro; IPR000403; PI3_PI4_kinase.
DR    InterPro; IPR008940; Prenyl_trans.
DR    Pfam; PF02259; FAT; 1.
DR    Pfam; PF02260; FATC; 1.
DR    Pfam; PF00454; PI3_PI4_kinase; 1.
DR    SMART; SM00146; PI3Kc; 1.
DR    PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
DR    PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR    PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR    PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW    Transferase; Kinase; Cell cycle; Repeat.
FT    REPEAT        1     31     HEAT 1.
FT    REPEAT      164    201     HEAT 2.
FT    REPEAT      331    371     HEAT 3.
FT    REPEAT      410    449     HEAT 4.
FT    REPEAT      474    512     HEAT 5.
FT    REPEAT      522    560     HEAT 6.
FT    REPEAT      562    596     HEAT 7.
FT    REPEAT      642    679     HEAT 8.
FT    REPEAT      684    722     HEAT 9.
FT    REPEAT      728    766     HEAT 10.
FT    REPEAT      843    880     HEAT 11.
FT    REPEAT      904    923     HEAT 12.
FT    REPEAT      924    961     HEAT 13.
FT    REPEAT      964   1003     HEAT 14.
FT    REPEAT     1005   1042     HEAT 15.
FT    REPEAT     1732   1769     HEAT 16.
FT    DOMAIN     1987   2335     PI3K/PI4K.
SQ    SEQUENCE   2335 AA;  266181 MW;  5DCF1CF4AEB8E9A4 CRC64;
```

```
Query Match       44.1%;  Score 4147;  DB 1;  Length 2335;
Best Local Similarity  45.9%;  Pred. No. 1.1e-246;
Matches  843;  Conservative 337;  Mismatches 516;  Indels 140;  Gaps  24;
```

```
Qy     1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDPNPGVINNVLA  60
          .:||:|:||||||:|| ||||:|  | ||||:||:||:||:||
Db   613  MEYSGNSRQKESAQLLKLLVSKARTIIKPYIQSIIHVILPKA---ADTSPGVSSAIIS  668

Qy    61  TIGELAQVSGLEMRKWVDELFI-IIMDMLQDSLLAKRQVALWTLGQLVASTGYVVEPYR  119
          :|||| |||||||||:||||:  |:||||||| | ||||| |  | :||||||||
Db   669  ALGELASVEGEDMPVDVRGSFMCLILVNLQDCSSTLKRLASLKCLRKLCGRSGYVIQPYL  728

Qy   120  KYPTLLEVLLNFLKTEQNGGTRREAIRVLGLLGALDPYKHHVKNIGMIDQSRDASAVSLSE  179
          |:| |||:|::|  :|||   :       ::|||||||:| |:||||
Db   729  DYPPLLGALIGILQSQKQPTPIRREVLRTLGVLGALDPYTY----------LTTEEVSD  776

Qy   180  SKSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVQAITFIF  239
          |:|:|||:||||||:||  ||||||||| |::|:||:|:||||||:||:||||: |:
Db   777  DLQSSHNNAHGVPQ--ISAAQYPSLENYAMVAVVTLIGILKDSSLSMHHSSVVQAVMHIC  834

Qy   240  KSLGLKCVQFLPQWMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMR  299
          |:|||||||:| ||||||:||  | ||| || ||||| :|||| :||:|||||:||||||
Db   835  SQMGSKSTVFLPQVVPTFLQVMQSLSASSAEFYFQQLTTILSIGPNIRDYVSDIFNLSK  894
```

```
Qy   300  EPWWNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPIHMLRVFMDNSPGRIVSIKLLAA  359
          :|: :| ||:||||| || :| ||  ||||:||||:|| || |||||| |:||
Db   895  VFWBSTTSLLIVILELIDAIAIALQDEFKFYLPQILSCMLKAFSLDN TSSRSVSYKVLQS  954

Qy   360  IQLFGANLDDYIHLLLPPIVKLPDAPAPLPSRKAALETVDRLTESLDPTDYASRIIHPI  419
          :||:|| | |::|| || ::|::|: |||::|   | || || ||||:|||||||||:
Db   955  FVIFGSNIEEYMHLVLPVIIRSFERDTIPLGFRKSALKCIAQLFQSVNFSDHASRIIHPL  1014

Qy   420  VRTLDQS-PELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVK  478
          || |:| :  |:|| || |:||:||| :|| |||:|||||||:||||| ||:| || :
Db  1015  VFMLGKSNGDLRAVIMDTLCAIVSQLGYDYSIFIPMVNKVLVSHKISHPAYELVLVSRLLK  1074

Qy   479  GYTLADEEDPLIYQHRMLRSGQGDALASGPVETGP-MKKLHVSTINLQKANGAARRVS  536
          |:| |:|: |:||:||:|| | |||||| || |  |||||:|||:|||| :||:
Db  1075  GEPL--PKDVVVKEFKPRPS------TKPFSTQDEVLTKLPVDQASLKAAWESSQKLT  1124

Qy   537  KDDWLEWLRRLSLELLKDSSPSLRSCWALAQAINPMARDLFNAAFVSCWSELNEDQQDE  596
          | | | || |:||||||||||| | |:|| |  |  || || ||| |:||||||:| ||
Db  1125  RDDWQDWIRRISIELLKESPSSALRSCSTLAGIYHPLARDLFNVSPLSCWDELTESNKKN  1184

Qy   597  LIRSIELALTSODIA-EVTQTLLNLAEFMEHSDKG-PLPLRDDNGIVLLGERAAKCRAYA  654
          |::|:| || |:    :| ||||:||||| |    |:|:| ||:|| | ||| | :
Db  1185  LVKSIELAMNAPNISVEILQTLLNLAEYMEREDHTLPIPIK------VISAHASKCNVYA  1238

Qy   655  KALHYKHLEF---QKGPTPAILESLSINNKLQQPRAAAGVLEYAMKHPGELEIQATWY  710
          ||| :| |:      ||| :|  |:|:|||: |:| | |:|||||| |||:|| |||
Db  1239  KALYTELQFVQETKEEVSISTIESLITINNHLQQSDAAVGMLQYTKEH-KQPSLKETWY  1297

Qy   711  EKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQA  770
          ||||:|||||:||||| |||:|| |||||:|||||||||||||:|:|| |:| ||:||:|
Db  1298  EKLHRWDDALAAYEHREREGDSSFENIGKLRCYYALGDWDHLSELAQKAWTISEQEHRE  1357

Qy   771  KMARMAARAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQCCIDKARDL  830
          |:|: ||||||||||||||:||||| |:|| |   :|||||| ||||||:|||| :|||||
Db  1358  AIAPLAAAAAWGLGQWNLISEYVSAMDRDPQDKFFSAISAVHLGQYNKAYGHIERIRDI  1417

Qy   831  LDABLTAMAGESYSRAYGAMVSCHMLSELEVIQYK---LNPERREIRQIWERLQGCQ  887
          | | : |  ||||||:|||:|  || |||||:| ||  :| |||:||||||:|||| |
Db  1418  LVNDLSSIIGESYNRAYGIMVKSQMLSELEEIIDYKKNMQYENNLDSLKKTWKRLEGCQ  1477

Qy   888  RIVEDWQKILMVRSLLVVSPHEDMRTWLKYASLCCKSGRLALAHKTLVLLLGVDPSRQLDH  947
          |::|||: ||| ||||: ||::   |:||:| || ||| |||: |:|| ||||||||: :|
Db  1478  KNVDVWHNTLRFRAIVLLSPQDSPEMWIKLADLCRRSDRLKLSNQCLTYLMGRDPSNA--Y  1535

Qy   948  PLPT---VHPQVTYAYMKMWKKSARKIDAFQHXQHFVQTMQQQAQHAIATEDQQH-----  999
          |:|:   || :||||:|| |||| :|| || |  :|:||| | | |: | |: |:
Db  1536  PLDSLKLLNPHVVTYLKYLNATDQKNIAVSELEEP-TSYLSSKHGYKMGDSSKLVDIL  1593

Qy  1000  -----KQELHKLMARCFLKLGEWQLNLQ-GINBSTIPKVLQYYSAATEHDRSWYKAWHAW  1053
               :|| ||:|||||||||||:|||:  |:|  |||:|:|||| | |||| |||:|||
Db  1594  ASSSVSSEERSFLARCFHKLGKWKKSLQDSVNQBSVRDILNCVFYATLFDKSWYKAWHSW  1653

Qy  1054  AVMNFRAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATTASTEGSNSESEA  1113
          |:||| |:|  |||
Db  1654  ALANFEVVGYY------------------------------------------  1664

Qy  1114  ESTENSPTPSPLQKKVTEDJSKTLMYTVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDY  1173
          |:|  |  |||::| :|||:|||||||||||| ::|:||:|||| :||||||||| |||
Db  1665  EQTEHG--------VTQDMYE---QYIVPAIKGFFHSSVLNQKNSLQDILRLLLNLWFX?  1712

Qy  1174  GHWPDVNEALVEGVKAIOIDTWLQVIPQLIARIDTPRPLVGRLIHQLLITDIGRYHPQALI  1233
          |  |:|| :|||||:| || |||:|||||:||:|  |:|:|||:||||:|| ||||||||:
Db  1713  GEHSDVAAAIVEGFSNVPMDTWLEVIPQSIARIHTSSSSVRASVHQLLSDIGRVHPQALV  1772

Qy  1234  YPLTVASKSTTTARHNAANKIIKNMCEHSNTLVQQAMMVSEELIRVAILWHEMWHEGLEE  1293
          |||:|||||:|:|  || | ||:|| | ||||| ::|:|::||||||:|| ||| |||||
Db  1773  YSLTVSSKSTNPQQKSAKSIMDSMLSHSDTLVRQALLVSQELIRVALLWHELWEYGLEE  1832

Qy  1294  ASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGN  1353
          ||:||||||||||||:||:|||||:|||||||||||||||||||||||::|||||:| |
Db  1833  ASQAYFSDHDISLMIDIVKPLHETELEKGPSTLSEISFAQTFGYDLRKARSTWQKFIQDGD  1892
```

Db 1114 ILRNAWEASQRSTKDDWQEWIRRLGVALLRESPSHALRACAALAAAYQPLARDLFNASFV 1173
Qy 584 SCWSELNEDQQDELIRSIELALTSQDIA-EVTQTLNIAEPMEHSDKGPLPLRDDNGIVL 642
Db 1174 SCWSELYDHFQEELVKSIEIALTSPHISPEIIQILLNLABPMEHDDK-PLPI----DIRT 1228
Qy 643 LGERAAKCRAYAKALHYKELEFQK-----GPTPAILESLISINNKLQQPEAAAGVLEYAM 697
Db 1229 LGAYAAKCHAPAKALHYKELEFIEEELVTKPSVDTIEALLISINNQLQQPDAAIGIIKHAQ 1288
Qy 698 KHFGELEIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEALGEWLGCLHQQC 757
Db 1289 QH-DKMNLKETWYEKLQRWEDALSAYEKREAAGAGNFBITMGKLRC.HALGEWDRLSQLA 1347
Qy 758 CEKWTLVNDETQAKMARMAAAAAAWGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLF 817
Db 1348 QENWIHAGHDARRYIAPLSVAAAWGLGWQEQMDEYISVMKSESPDKAFFNAIVALHRSQF 1407
Qy 818 SLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEBVIQYKLV--PERRRI 874
Db 1408 BEAASYITRARDLLDTELITALVGESYNRAYSVAVRVQMLSELEBIITYKKAEDKPEVRRM 1467
Qy 875 IRQIWERLGQGCQRIVEDWQKILMYRSLVVSPHEDMRTWLKYASLCGKSGRLALAHKILV 934
Db 1468 IKKTWVRRLRKGCQRNVDVWQRMLRIRSLVISPRDNMEMWTKFANLCRKSGRISLAKKSLN 1527
Qy 935 LLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQEMQHFVQTMQQ-------- 986
Db 1528 LLLEDDENLLDNSLVLKKTHPSIVLKPLWAUDDKRKALNSMQEFTSQLLISDINVDPAL 1587
Qy 987 --QAQHAIATEDQQHKQELHKIMARCFLKLGEWQLNLQ-GINESTIPKVLQYYSAATEHD 1043
Db 1588 FVQSTSVNTQKSQEBIQYYFHLLARCYHKQGGWQQEIENNWSEGSFDGVLQSYMYATQFD 1647
Qy 1044 RSWYKAWHAWAVMNPEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTAATATTTAS 1103
Db 1648 SKWYKAWHSWALANPEAVKFLE------------------------------------ 1669
Qy 1104 TEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTPAVQGFFRSISLSRGNNLQDT 1163
Db 1670 -----------QSEEQIPSAAYEQ------------YIIPAVKGFFKSIALSKG-NLQDT 1705
Qy 1164 LRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTFRPLVGRLIHQLLTD 1223
Db 1706 LRLLNLWFKFGKNNSNVINTLNVGISTVNIDIWLDVIPQLIARIHASSLSVRKSVHQLLSD 1765
Qy 1224 IGRYHPQALLYPLTVASKSTTARHNAANKIIKNMCEHSNTLVQQAMMVSBELIRVAILW 1283
Db 1766 VGRAHPQALLYPLTVAAKSQSSARQNAALAIMDSLKTHSPRLVEQARLVSHELIRAAILW 1825
Qy 1284 HEMWHBGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE 1343
Db 1826 HEQWEHGLEEASRLYFGDHNIEGMFAVLRPLHEMLERGPETLREISFQQAFGRDLVEARD 1885
Qy 1344 WCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLEBLQYVSPKLLMCRDLELAVPGT 1403
Db 1886 CCIRFEQTGDISDLNQAWDLYYQVFKKIRKQLPQLTTLDLQYVSPKLLHVHDLELIAVPGT 1945
Qy 1404 YDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFVFLLKGHEDLRQDERVMQLFG 1463
Db 1946 YVSGKPVIRIVKFYPTFNVITSKQRPRRJSIKGSDGKDYQYVLKGHEDIRQDERVMQLFG 2005
Qy 1464 LVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWPHCDTLHALIRDYREKKKILLNI 1523
Db 2006 LCNNLLLADPETFKRLLSIQRYPVIPLSPDSGLLGWVLDSDTLHVLIRDYRESRKILLNI 2065
Qy 1524 EHRIMLRMAPDYDHLILMQKVEHVEHAVNNTAGDIAKLIWLKSPSSEVWFDRRTNYTRS 1583
Db 2066 EHRLIIQMADYDRLTLLQKVEVFEYALLSTTGQDIYRVIWLKSRSSEAWLNRRTNYSRT 2125
Qy 1584 LAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGDCFEVAMTREKFPEKIPFRLTRML 1643
Db 2126 LAVMSMVGYILGLGDRHPSNLMLDRYTGNIIHIDFGDCFEVAMHREKFPEKIPFRLTRML 2185

Qy 1644 TNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEAFVYDPLINWRLMDINTKGNKESR 1703
Db 2186 VNAMEVSGIEGTFRITCEHVMRVLRTNKESVMAVLEAFVYDPLINWRLAPAYSPSIDEKQ 2245
Qy 1704 TRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIHSFTGDGLVKPEPALNKKAIQIINR 1763
Db 2246 SNEPNTLLGBTIDGLHRKRLNE-----EGTTLEER-------QKPEIINQRAITTVLNR 2291
Qy 1764 VRDKLTGRDFSHDDTLDVPTQVELLIKQATSHBNLCQCYIGWCPFW 1809
Db 2292 VSNKLTGRDFKPQQQLDVPSQVEKLILQATSIENLCLCYIGWCSFW 2337

RESULT 5
TOR1_SCHPO
ID TOR1_SCHPO STANDARD; PRT; 2335 AA.
AC O14356;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol 3-kinase tor1 (EC 2.7.1.137) (PI3-kinase)
DE (PtdIns-3-kinase) (PI3K)
GN TOR1 OR SPBC30D10.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP IDENTIFICATION, FUNCTION, AND INDUCTION.
RX MEDLINE=21269225; PubMed=11096119;
RA Weisman R., Choder M.;
RT "The fission yeast TOR homolog, tor1+, is required for the response to
RT starvation and other stresses via a conserved serine.";
RL J. Biol. Chem. 276:7027-7032(2001).
CC -!- FUNCTION: Phosphatidylinositol 3-kinase homolog required for G1
CC progression and entry into stationary phase. Also required for the
CC onset of meiosis and sporulation under nitrogen and carbon
CC starvation conditions.
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC 1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC -!- INDUCTION: By nitrogen and/or carbon starvation, cold, osmotic and

```
ID   TOR2_SCHPO     STANDARD;      PRT;   2337 AA.
AC   Q9Y7K2; Q94507;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Phosphatidylinositol 3-kinase tor2 (EC 2.7.1.137) (PI3-kinase)
DE   (PtdIns-3-kinase) (PI3K).
GN   TOR2 OR SPBC216.07C OR SPBC646.01C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
RN   [2]
RP   IDENTIFICATION.
RX   MEDLINE=21269225; PubMed=11096119;
RA   Weisman R., Choder M.;
RT   "The fission yeast TOR homolog, tor1, is required for the response to
RT   starvation and other stresses via a conserved serine.";
RL   J. Biol. Chem. 276:7027-7032(2001).
CC   -!- FUNCTION: Phosphatidylinositol 3-kinase homolog required for G1
CC       progression (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC       1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   -!- SIMILARITY: Contains 15 HEAT repeats.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AL035216; CAA22805.1; -.
DR   EMBL; AL049558; CAB40167.1; -.
DR   PIR; T39913; T40577.
DR   HSSP; P42345; 1AUE.
DR   GeneDB_SPombe; SPBC216.07c; -.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000357; HEAT.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS50077; HEAT_REPEAT; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; FALSE_NEG.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Transferase; Kinase; Cell cycle; Repeat.
FT   REPEAT    153    200   HEAT 1.
FT   REPEAT    249    286   HEAT 2.
FT   REPEAT    290    326   HEAT 3.
FT   REPEAT    409    446   HEAT 4.
FT   REPEAT    474    512   HEAT 5.
FT   REPEAT    559    596   HEAT 6.
FT   REPEAT    642    679   HEAT 7.
FT   REPEAT    683    721   HEAT 8.
FT   REPEAT    727    765   HEAT 9.
FT   REPEAT    802    840   HEAT 10.
FT   REPEAT    844    880   HEAT 11.
FT   REPEAT    881    921   HEAT 12.
FT   REPEAT    965   1004   HEAT 13.
FT   REPEAT   1006   1043   HEAT 14.
FT   REPEAT   1735   1772   HEAT 15.
FT   DOMAIN   1990   2337   PI3/PI4K.
SQ   SEQUENCE   2337 AA;  266376 MW;  190F448DAC4FD2D9 CRC64;

Query Match          46.7%;    Score 4398;  DB 1;  Length 2337;
Best Local Similarity 48.1%;   Pred. No. 3.9e-262;
Matches 888;  Conservative 324;  Mismatches 476;  Indels 158;  Gaps 20;

Qy     1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKDPPDENPGVINNVLA   60
           |::||:|::|  ||::|::|::|||  |::||||  :|::||::|::|   |:|
Db   613 LDYSTIIRTKEENAKLLCLLIAAAPRLIESHVEPILQILLPKAKDSSSIVAASIVN---  668

Qy    61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVEPYRK  120
           ::|  |:||||:|  | :::|  ::| | |:||:: |: |:|| ||  :||||:||:|
Db   669 SLGEICQISGEVIVPPIKDLMPLIEALQDQSSPIRRAAALKALGNLSSSTGVVIDPYIE  728

Qy   121 YPTLLEVLINFLKTEQNQGTRRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
           :|:||::||  |:|||  | :|||::|:|| |||||||: |:|
Db   729 FPSLLDILIGITKTEQDITIRRETIKLIGTLGALDPNRHRV                     769

Qy   181 KSSQQSSDYSTSEMLVNMGNLP-----------------LDEPYPAVSMVALMRIFRDQSLSHH 227
           |::|   ::::|::|::|                     |||:|::|:|::|::| |::|::|
Db   770 -----LEKGTEKVVPEQKNIPPDISLLMSGIGPSSDEYYPIVVITALMSILKDPSLJTIH    823

Qy   228 HTMVVQAITFIFKSLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHI  287
           ||:|:||::|||::||||| ||||::||:|::|:|||  ::||||:||||::|:|::::|||
Db   824 HTAVIQAVMYIFKTMGLRCAPFLSQIIPEFIAVMRTCPTNILEFYFQQLSILVLIVRQHI  883

Qy   288 RPYMDEIVTLMREFWMNTSIQSTILLLIEQIVALGGEFKLYLPOLIPHMLRVFMEDNS   347
           |  ::|::||:::||||  :: ||: |:|::|:||:|||| |:|:||:::|:|:::|:|
Db   884 RSFLPDLFKLIKDFWNPHSNLQFTILSLIBSLARAMQGEFKPYLPSLLVVMLQIFPSDVS  943

Qy   348 PGRIVSIKLLAAIQLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLD  407
           :|::|::|:|:|:|::||::|||:|::||||::|::|  ::| |:|:: |:||:||::|::
Db   944 VDSVSTKKVLHAFIVFGDTLADYFHMLLDPILRLYERNDVSIGIKESIMITIGRLSMVIN 1003

Qy   408 FTDYASRIIHPIVRTLD-QSPELRSTAMDTLSSIVFQLGKKYQIPIPMVNKVLVRHRINH  466
           ::::|||||||:::.| |    ::| ::||:|:|:||:: |::::|:||:|:|:|:||:|:
Db  1004 LSEYASRIIHPVMRVLSCNNASLIRVSMDTVCALIYQLNVDFAIFIPMDKCLKMNGVTH  1063

Qy   467 QRYDVLICRIVKGYTLADEEBDL---IYCHRMLRSGQGDLALASGPVETGPMKKLHVSTI  523
             ::|::|::||:|::|::::|         ::|::  ||:|::|::::|::|:|||:|::
Db  1064 ETYSILVEQFL------QEQLPIKLNPYEKYDKPLDVVASAADITS--KKLPVNQE     1113

Qy   524 NLQKAWGAARVSKDDWLEWLRRLSLELLKDSSSPSLRSCWALIAQAYNPMARDLFNAAFV  583
           |:::|::|||:||||||:|||||:::|:|:|:||:|:|::|||  |||||:|:|:|||::|
```

```
DR    PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR    PROSITE; PS50290; PI3_4_KINASE_3; 1.
DR    PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
KW    Transferase; Kinase; Repeat.
FT    REPEAT       16      53       HEAT 1.
FT    REPEAT      650     688       HEAT 2.
FT    REPEAT      859     897       HEAT 3.
FT    REPEAT      988    1025       HEAT 4.
FT    REPEAT     1069    1106       HEAT 5.
FT    REPEAT     1109    1148       HEAT 6.
FT    REPEAT     1150    1186       HEAT 7.
FT    DOMAIN     1382    1982       FAT.
FT    REPEAT     1933    1970       HEAT 8.
FT    DOMAIN     2182    2549       PI3K/PI4K.
SQ    SEQUENCE   2549 AA;  288791 MW;  BB41EA7B9086F99 CRC64;

Query Match             99.2%;  Score 9341;  DB 1;  Length 2549;
Best Local Similarity   99.1%;  Pred. No. 0;
Matches 1793;  Conservative 5;  Mismatches 11;  Indels  0;  Gaps  0;

Qy          1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA  60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDPDNPGVINNVLA  800

Qy         61  TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVEPYRK  120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801  TIGELAQVSGLEMRKWVDELFVIIMDMLQDSSLLAKRQVALWTLGQLVASTGVVEPYRK  860

Qy        121  YPTILEVLLNFLKTEQNGGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  180
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861  YPTILEVLLNFLKTEQNQGTRRRAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  920

Qy        181  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK  240
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK  980

Qy        241  SIGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE  300
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981  SIGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE  1040

Qy        301  FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI  360
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041  FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSQRIVSIKLLAAI  1100

Qy        361  QLFGANLDDYIHILLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV  420
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101  QLFGANLDDYIHILLPPIVKLFDAPEVPLPSRKAALETVDRLTESLDFTDYASRIIHPIV  1160

Qy        421  RTLDQSPELRSTAMDTLSSLIVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY  480
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161  RTLDQSPELRSTAMDTLSSLIVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY  1220

Qy        481  TLADEEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW  540
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1221  TLADEEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW  1280

Qy        541  LEWLRRLSLELLKDSSSPPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  600
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1281  LEWLRRLSLELLKDSSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  1340

Qy        601  IELALTSQDIAEVTCYTLLNLARFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK  660
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1341  IELALTSQDIAEVTCYTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK  1400

Qy        661  ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL  720
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1401  ELEFQKGPTPAILESLISINNKLQQPEAASGVLEYAMKHFGELEIQATWYEKLHEWEDAL  1460

Qy        721  VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA  780
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1461  VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA  1520

Qy        781  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG  840
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1521  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG  1580

Qy        841  ESYSRAYGAMVSCHMLSELEEVIQYKLVPERRELIRQIWWERLQGCQRIVEDWQKILMVR  900
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1581  ESYSRAYGAMVSCHMLSELEEVIQYKLVPERRELIRQIWWERLQGCQRIVEDWQKILMVR  1640

Qy        901  SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY  960
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1641  SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY  1700

Qy        961  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN  1020
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1701  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKIMARCFIKLGEWQLN  1760

Qy       1021  LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKQNQARDEKKKLRHA  1080
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1761  LQGINESTIPKVLQYYSAATEHDRSWTKAWHAWAVMNFEAVLHYKQNQARDEKKKLRHA  1820

Qy       1081  SGANITNATTAATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY  1140
               |||||||||||||||||||| ::||||||||||||||||||||||||||||||||||||
Db       1821  SGANITNATTTATTAASAAAATSTEGSNESEAESNSSPTPSPLQKKVTEDLSKTLLLY  1880

Qy       1141  TVPAVQGFFRSISLSRGNNLQDTLRVLTLMFDYGHWPDVNEALVEGVKAIQIDTWLQVIP  1200
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1881  TVPAVQGFFRSISLSRGNNLQDTLRVLTLMFDYGHWPDVNEALVEGVKAIQIDTWLQVIP  1940

Qy       1201  QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALLYPLTVASKSTTTARHNAANKILKNMCE  1260
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1941  QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALLYPLTVASKSTTTARHNAANKILKNMCE  2000

Qy       1261  HSNTLVQQAMMVSEELIRVALLWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER  1320
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001  HSNTLVQQAMMVSEELIRVALLWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER  2060

Qy       1321  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS  1380
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS  2120

Qy       1381  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTMGSNGH  1440
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTMGSNGH  2180

Qy       1441  EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV  1500
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181  EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV  2240

Qy       1501  PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA  1560
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241  PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA  2300

Qy       1561  KLIWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD  1620
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2301  KLIWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD  2360

Qy       1621  CFEVAMTREKFPPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  1680
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2361  CFEVAMTREKFPPEKIPFRLTRMLTNAMEVTGLDGNYRTTCHTVMEVLREHKDSVMAVLEA  2420

Qy       1681  FVYDPLLNWRLMDTNTKGNKKSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  1740
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2421  FVYDPLLNWRLMDTNAKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  2480

Qy       1741  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  1800
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2481  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  2540

Qy       1801  CYIGWCPFW  1809
               |||||||||
Db       2541  CYIGWCPFW  2549
```

RESULT 4

```
Db 1281 LEWLRLSLELLLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2361 CFEVAMTREKFPEKIPFRLITRMLTNAMEVTGLDGNYRTTCHIVMEVLREHKDSVMAVLEA 2420

Qy  601 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNG?VLLGERAAKCRAYAKALHYK 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1681 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKGTTVPESIH 1740

Db 1341 IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2421 FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKAGTTVPESIH 2480

Qy  661 ELEFQKGPTPAILESLSISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYKLHEWEDAL 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1741 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 1800

Db 1401 ELEFQKGPTPAILESLSISINNKLQQPEAASGVLEYAMKHFGELEIQATWYKLHEWEDAL 1460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2481 SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ 2540

Qy  721 VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1801 CYIGWCPFW 1809

Db 1461 VAYDKKMDTNKEDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2541 CYIGWCPFW 2549

Qy  781 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1521 WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy  841 ESYSRAYGAMVSCHMLSELEVIQYKLVPEREEIIRQIWWERLQGCCGRIVEDWQKILMVR 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1581 ESYSRAYGAMVSCHMLSELEVIQYKLVPEREEIIRQIWWERLQGCQRIVEDWQKILMVR 1640

Qy  901 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTTYAY 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1641 SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTAHPQVTTYAY 1700

Qy  961 MKNMWKSARKIDAFQHEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1701 MKNMWKSARKIDAFQHEMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN 1760

Qy 1021 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1761 LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy 1081 SGANITNATTAATTAATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||:|::||||||||
Db 1821 SGANITNATTAATTAASAAAATSTEGSNSESEAESNENSPTPSPLQKKVTEDLSKTLLLY 1880

Qy 1141 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1881 TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy 1201 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNNAANKILKNMCE 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1941 QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNNAANKILKNMCE 2000

Qy 1261 HSNTIVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2001 HSNTIVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER 2060

Qy 1321 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPCLTS 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2061 GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPCLTS 2120

Qy 1381 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTMGSNGH 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2121 LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTMGSNGH 2180

Qy 1441 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2181 EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV 2240

Qy 1501 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2241 PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA 2300

Qy 1561 KLLWIKSPSSEVWFDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMLDRLSGKILHIDFGD 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2301 KLLWIKSPSSEVWFDRRTNYTRSLAVMSMVGYIIGLGDRHPSNLMLDRLSGKILHIDFGD 2360

Qy 1621 CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA 1680
```

RESULT 3

```
FRAP_RAT
ID   FRAP_RAT       STANDARD;      PRT;   2549 AA.
AC   P42346;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   FKBP-rapamycin associated protein (FRAP) (Rapamycin target protein).
GN   FRAP1 OR RAFT1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1],
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95122530; PubMed=7822316;
RA   Sabers C.J., Martin M.M., Brunn G.J., Williams J.M., Dumont P.J.,
RA   Wiederrecht G., Abraham R.T.;
RT   "Isolation of a protein target of the FKBP12-rapamycin complex in
RT   mammalian cells.";
RL   J. Biol. Chem. 270:815-822(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=94306515; PubMed=7518356;
RA   Sabatini D.M., Erdjument-Bromage H., Lui M., Tempst P.,
RA   Snyder S.H.;
RT   "RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-
RT   dependent fashion and is homologous to yeast TORs.";
RL   Cell 78:35-43(1994).
CC   -!- FUNCTION: ACTS AS THE TARGET FOR THE CELL-CYCLE ARREST AND
CC       IMMUNOSUPPRESSIVE EFFECTS OF THE FKBP12-RAPAMYCIN COMPLEX.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   -!- SIMILARITY: Contains 8 HEAT repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L37085; AAA65929.1; -.
DR   EMBL; U11681; AAA20091.1; -.
DR   PIR; A54837; A54837.
DR   HSSP; P42345; 1FAP.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000357; HEAT.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008940; Prenyl_trans.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
```

```
Qy  1261  HSNTLVQQAMMVSEELIRVAIIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2001  HSNTLVQQAMMVSEELIRVAIIWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER  2060

Qy  1321  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLITQAWDLYYHVFRRISKQLPQLTS  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2061  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLITQAWDLYYHVFRRISKQLPQLTS  2120

Qy  1381  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2121  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH  2180

Qy  1441  EFVPLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRRNLSIQRYAVIPLSTNSGLIGWV  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2181  EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRRNLSIQRYAVIPLSTNSGLIGWV  2240

Qy  1501  PHCDTLHALIRDYRERKKKLLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2241  PHCDTLHALIRDYRRKKKLLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDLA  2300

Qy  1561  KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2301  KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD  2360

Qy  1621  CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2361  CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  2420

Qy  1681  FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2421  FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  2480

Qy  1741  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2481  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  2540

Qy  1801  CYIGWCPFW  1809
          |||||||||
Db  2541  CYIGWCPFW  2549
```

RESULT 2
```
FRAP_MOUSE
ID   FRAP_MOUSE      STANDARD;      PRT;    2549 AA.
AC   Q9JLN9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   FKBP-rapamycin associated protein (FRAP).
GN   FRAP1 OR FRAP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Bliskovsky V., Mock B.;
RT   "Positional cloning of mouse plasmacytoma susceptibility gene.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ACTS AS THE TARGET FOR THE CELL-CYCLE ARREST AND
CC       IMMUNOSUPPRESSIVE EFFECTS OF THE FKBP12-RAPAMYCIN COMPLEX.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   -!- SIMILARITY: Contains 8 HEAT repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   --------------------------------------------------------------------------
DR   EMBL; AF152838; AAF73196.1; -.
DR   HSSP; P42345; 1FAP.
DR   MGD; MGI:1928394; Frap1.
DR   GO; GO:0007281; P:germ-cell development; IDA.
DR   InterPro; IPR008838; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000357; HEAT.
DR   InterPro; IPR000403; PI3_PI4 kinase.
DR   InterPro; IPR008940; Prenyl_trans.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
DR   PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
KW   Transferase; Kinase; Repeat.
FT   REPEAT       16     53       HEAT 1.
FT   REPEAT      650    688       HEAT 2.
FT   REPEAT      859    897       HEAT 3.
FT   REPEAT      988   1025       HEAT 4.
FT   REPEAT     1069   1106       HEAT 5.
FT   REPEAT     1109   1148       HEAT 6.
FT   REPEAT     1150   1186       HEAT 7.
FT   DOMAIN     1382   1982       FAT.
FT   REPEAT     1933   1970       HEAT 8.
FT   DOMAIN     2182   2549       PI3K/PI4K.
SQ   SEQUENCE   2549 AA;  288734 MW;  C82EB36EF861BBC3 CRC64;
```

```
Query Match             99.3%;  Score 9343;  DB 1;  Length 2549;
Best Local Similarity   99.2%;  Pred. No. 0;
Matches 1794;  Conservative    4;  Mismatches    4;  Indels   11;  Gaps   0;
```

```
Qy     1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDDPDNPGVINNVLA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKLKDPDDPDNPGVINNVLA   800

Qy    61  TIGELAQVSGLEMRKWTVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801  TIGELAQVSGLEMRKWTVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK   860

Qy   121  YPTLLEVLLNFLKTEBQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861  YPTLLEVLLNFLKTEBQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  920

Qy   181  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITFIFK   980

Qy   241  SIGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQOLGMLVSFVKSHIRPYMDEIVTLMRE   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981  SIGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQOLGMLVSFVKSHIRPYMDEIVTLMRE  1040

Qy   301  FWMMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1041  FWMMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMEDNSQGRIVSIKLLAAI  1100

Qy   361  QLFGANLDDYLIHLLLPPIVKJFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1101  QLFGANLDDYLIHLLLPPIVKJFDAPEVELPSRKAALETVDRLTESLDFTDYASRIIHPIV  1160

Qy   421  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKLVLVRRHINHQRYDLICRIVKGY   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1161  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKLVLVRRHINHQRYDLICRIVKGY  1220

Qy   481  TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1221  TLADEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW  1280

Qy   541  LEWLRRLSLELLKLKDSSPSLRSCWALAQAYNPMARDLFNAAPVSCWSELNEDQQDELIRS  600
```

```
DR   EMBL; L34075; AAA58486.1; -.
DR   EMBL; U88966; AAC39933.1; -.
DR   PIR; S45340; S45340.
DR   PDB; 1FAP; 23-JUL-97.
DR   PDB; 1NSG; 18-MAR-98.
DR   PDB; 1AUE; 18-NOV-98.
DR   PDB; 2FAP; 09-AUG-99.
DR   PDB; 3FAP; 13-SEP-00.
DR   PDB; 4FAP; 13-SEP-00.
DR   Genew; HGNC:3942; FRAP1.
DR   MIM; 601231; -.
DR   GO; GO:0000074; P:regulation of cell cycle; TAS.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000357; HEAT.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR008940; Prenyl_trans.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
DR   PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
KW   Transferase; Kinase; Repeat; 3D-structure.
FT   REPEAT          16     53       HEAT 1.
FT   REPEAT         650    688       HEAT 2.
FT   REPEAT         859    897       HEAT 3.
FT   REPEAT         988   1025       HEAT 4.
FT   REPEAT        1069   1106       HEAT 5.
FT   REPEAT        1109   1148       HEAT 6.
FT   REPEAT        1150   1186       HEAT 7.
FT   DOMAIN        1382   1982       FAT.
FT   REPEAT        1933   1970       HEAT 8.
FT   DOMAIN        2182   2549       PI3K/PI4K.
FT   CONFLICT       353    353       K -> N (IN REF. 2).
FT   CONFLICT       359    359       S -> N (IN REF. 2).
FT   CONFLICT       364    364       D -> N (IN REF. 2).
FT   CONFLICT       390    390       M -> L (IN REF. 2).
FT   CONFLICT       430    430       R -> L (IN REF. 2).
FT   CONFLICT       455    457       VLD -> GVE (IN REF. 2).
FT   CONFLICT       461    461       A -> G (IN REF. 2).
FT   CONFLICT       482    484       VFT -> FFN (IN REF. 2).
FT   CONFLICT       489    489       L -> V (IN REF. 2).
FT   CONFLICT       513    513       L -> I (IN REF. 2).
FT   CONFLICT       539    539       L -> V (IN REF. 2).
FT   CONFLICT       553    553       R -> C (IN REF. 2).
FT   CONFLICT       956    999       MRIFRDQSLSHHTMVVQAITPIFKSLGLKCVQFLPQVMPT
FT                                   FLN -> ADLPRPVTLSSSHHGCPGHHLHLQVPGTQMCAV?
FT                                   APGHAHVPY (IN REF. 2).
FT   CONFLICT      1075   1075       I -> S (IN REF. 2).
FT   HELIX         2023   2041
FT   HELIX         2044   2060
FT   HELIX         2065   2091
FT   HELIX         2094   2111
SQ   SEQUENCE   2549 AA;  288888 MW;  7D9AD6784882AB4 CRC64;

Query Match             100.0%;  Score 9413;  DB 1;  Length 2549;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1809;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKKDPPDDNPGVINNVLA   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741  LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKKDPPDDNPGVINNVLA  800

Qy     61  TIGELAQVSGLEMRKKVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801  TIGELAQVSGLEMRKKVDELFIIIMDMLQDSSLLAKRQVALWTLGQLVASTGYVVEPYRK  860

Qy    121  YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    861  YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  920

Qy    181  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITPIFK  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    921  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHTMVVQAITPIFK  980

Db    241  SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    981  SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy    301  FWVWNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041  FWVWNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 1100

Qy    361  QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101  QLFGANLDDYLFLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160

Qy    421  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINEQRYDVLICRIVKGY  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINEQRYDVLICRIVKGY 1220

Qy    481  TLADEEBDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221  TLADEEBDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW 1280

Qy    541  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS 1340

Qy    601  IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341  IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK 1400

Qy    661  ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401  ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL 1460

Qy    721  VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461  VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA 1520

Qy    781  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALEQDLFSLAQQCIDKARDLLDAELTAMAG  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG 1580

Qy    841  ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581  ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCQRIVEDWQKILMVR 1640

Qy    901  SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641  SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLGVDPSRQLDHPLPTVHPQVTYAY 1700

Qy    961  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKIMARCFLKLGEWQLN 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKIMARCFLKLGEWQLN 1760

Qy   1021  LQGINESTIPKVLQYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761  LQGINESTIPKVLQYSAATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHA 1820

Qy   1081  SGANITNATTIAATTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1821  SGANITNATTIAATTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMY 1880

Qy   1141  TVPAVQGFFRSISLSRGNNLQDTLRVELMFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881  TVPAVQGFFRSISLSRGNNLQDTLRVILMFDYGHWPDVNEALVEGVKAIQIDTWLQVIP 1940

Qy   1201  QLIARIDTPRPLVGHLIHQLLTDIGRYHPQALIYPLIVASKSTTARHNAANKILRNMCE 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941  QLIARIDTPRPLVGHLIHQLLTDIGRYHPQALIYPLIVASKSTTARHNAANKILRNMCE 2000
```

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     March 2, 2004, 19:55:22 ; Search time 25 Seconds
                                    (without alignments)
                                    3767.796 Million cell updates/sec

Title:          US-09-517-491-12
Perfect score:  9413
Sequence:       1 LEHSGIGRIKEQSARMLGHL.........KQATSHENLCQCYIGWCPFW 1809

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters:     141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    SwissProt_42:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9413	100.0	2549	1	FRAP_HUMAN	P42345 homo sapien
2	9343	99.3	2549	1	FRAP_MOUSE	Q9JLN9 mus musculu
3	9341	99.2	2549	1	FRAP_RAT	P42346 rattus norv
4	4398	46.7	2337	1	TOR2_SCHPO	Q9Y7K2 schizosacch
5	4147	44.1	2335	1	TOR1_SCHPO	O14356 schizosacch
6	4105	43.6	2473	1	TOR2_YEAST	P32600 saccharomyc
7	3962	42.1	2470	1	TOR1_YEAST	P35169 saccharomyc
8	725.5	7.7	2386	1	RAD3_SCHPO	Q02099 schizosacch
9	631	6.7	3056	1	ATM_HUMAN	Q13315 homo sapien
10	628.5	6.7	3066	1	ATM_MOUSE	Q62388 mus musculu
11	628	6.7	2368	1	ESR1_YEAST	P38111 saccharomyc
12	564	6.0	4128	1	PRKD_HUMAN	P78527 homo sapien
13	552.5	5.9	4128	1	PRKD_MOUSE	P97313 mus musculu
14	494	5.2	2787	1	TEL1_YEAST	P38110 saccharomyc
15	390.5	4.1	3655	1	YAMB_SCHPO	Q10064 schizosacch
16	387	4.1	2565	1	TRAP_MOUSE	Q80YV3 mus musculu
17	380	4.0	3859	1	TRAP_HUMAN	Q9Y4A5 homo sapien
18	370	3.9	3744	1	TRA1_YEAST	P38111 saccharomyc
19	339	3.6	3803	1	TRA1_DROME	Q8IBU7 drosophila
20	226	2.4	875	1	VP34_YEAST	P22543 saccharomyc
21	219	2.3	801	1	VP34_SCHPO	P50520 schizosacch
22	216	2.3	1043	1	P11D_MOUSE	O35904 mus musculu
23	212.5	2.3	812	1	P3K3_SOYBN	P42348 glycine max
24	211.5	2.2	1020	1	VP34_CANAL	Q92213 candida alb
25	205.5	2.2	814	1	P3K1_SOYBN	P42347 glycine max
26	205	2.2	814	1	PI3K_ARATH	P42339 arabidopsis
27	204.5	2.2	1093	1	PI4K_DICDI	P54677 dictyosteli
28	201.5	2.1	1044	1	P11D_HUMAN	O00329 homo sapien
29	201.5	2.1	1068	1	P11A_MOUSE	P42337 mus musculu
30	201.5	2.1	1755	1	PEPL_MOUSE	Q9z269 mus musculu
31	196.5	2.1	1068	1	P11A_BOVIN	P32871 bos taurus
32	196.5	2.1	1068	1	P11A_HUMAN	P42336 homo sapien
33	195	2.1	1448	1	PK3G_HUMAN	O75747 homo sapien
34	192.5	2.0	2044	1	PI4K_HUMAN	P42356 homo sapien
35	192.5	2.0	3660	1	DMD_CHICK	P11533 gallus gall
36	190.5	2.0	1146	1	AGE1_CAEEL	Q94125 caenorhabdi
37	188	2.0	1505	1	PK3G_RAT	O70173 rattus norv
38	186.5	2.0	816	1	P3K4_DICDI	P54676 dictyosteli
39	186	2.0	2364	1	SPCO_HUMAN	Q01082 homo sapien
40	185.5	2.0	1102	1	P11G_MOUSE	Q9jhg7 mus musculu
41	183.5	1.9	1506	1	PK3G_MOUSE	O70167 mus musculu
42	183	1.9	1101	1	P11G_HUMAN	P48736 homo sapien
43	182.5	1.9	8797	1	SNE1_HUMAN	Q8nf91 homo sapien
44	181.5	1.9	1070	1	P11B_RAT	Q92110 rattus norv
45	178	1.9	851	1	YDBG_SCHPO	Q10366 schizosacch

```
                              ALIGNMENTS

RESULT 1
FRAP_HUMAN
ID   FRAP_HUMAN     STANDARD;      PRT;   2549 AA.
AC   P42345; Q9Y4I3;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   FKBP-rapamycin associated protein (FRAP) (Rapamycin target protein).
DE   FRAP1 OR FRAP OR FRAP2.
GN   Homo sapiens (Human).
OS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC   NCBI_TaxID=9606;
OX   [1]
RN   SEQUENCE FROM N.A.
RP   TISSUE=Brain;
RC   MEDLINE=94277209; PubMed=8008069;
RX   Brown E.J., Albers M.W., Shin T.B., Ichikawa K., Keith C.T.,
RA   Lane W.S., Schreiber S.L.;
RA   "A mammalian protein targeted by G1-arresting rapamycin-receptor
RT   complex.";
RT   Nature 369:756-758(1994).
RL   [2]
RN   SEQUENCE FROM N.A.
RP   MEDLINE=98317532; PubMed=9653645;
RX   Onyango P., Lubyova B., Gardellin P., Kurzbauer R., Weith A.;
RA   "Molecular cloning and expression analysis of five novel genes in
RT   chromosome 1p36.";
RT   Genomics 50:187-198(1998).
RL   [3]
RN   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 2018-2112.
RP   MEDLINE=96279639; PubMed=8662507;
RX   Choi J., Chen J., Schreiber S.L., Clardy J.;
RA   "Structure of the FKBP12-rapamycin complex interacting with the
RT   binding domain of human FRAP.";
RT   Science 273:239-242(1996).
RL   [4]
RN   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 2018-2112.
RP   MEDLINE=99190960; PubMed=10089303;
RX   Liang J., Choi J., Clardy J.;
RA   "Refined structure of the FKBP12-rapamycin-FRB ternary complex at 2.2
RT   A resolution.";
RT   Acta Crystallogr. D 55:736-744(1999).
RL   -!- FUNCTION: ACTS AS THE TARGST FOR THE CELL-CYCLE ARREST AND
CC       IMMUNOSUPPRESSIVE EFFECTS OF THE FKBP12-RAPAMYCIN COMPLEX.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   -!- SIMILARITY: Contains 8 HEAT repeats.
CC   -------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

30	227.5	873	2	T25442	hypothetical prote
31	226	875	2.4	A36369	1-phosphatidylinos
32	219	801	2.3	T52538	1-phosphatidylinos
33	216	664	2.3	PC4002	phosphatidylinosit
34	216	1043	2.3	T43502	1-phosphatidylinos
35	212.5	812	2.3	T07745	phosphatidylinosit
36	212.5	1876	2.3	T13801	phospboinositide 3
37	211.5	1020	2.2	T18260	1-phosphatidylinos
38	205.5	814	2.2	T07761	phosphatidylinosit
39	205	814	2.2	B96630	Phosphatidylinosit
40	204.5	1093	2.2	T18275	1-phosphatidylinos
41	200	3944	2.2	T19997	hypothetical prote
42	199.5	732	2.1	T08420	1-phosphatidylinos
43	199.5	2121	2.1	T27406	1-phosphatidylinos
44	196.5	1068	2.1	A43322	hypothetical prote
45	196.5	1068	2.1	I38110	1-phosphatidylinos

ALIGNMENTS

RESULT 1
S45340
FKBP-rapamycin-associated protein (FRAP) - human
C;Species: Homo sapiens (man)
C;Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C;Accession: S45340
R;Brown, E.J.; Albers, M.W.; Shin, T.B.; Ichikawa, K.; Keith, C.T.; Lane, W.S.; Schreib
Nature 369, 756-758, 1994
A;Title: A mammalian protein targeted by G1-arresting rapamycin-receptor complex.
A;Reference number: S45340; MUID:94277209; PMID:8008069
A;Accession: S45340
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2549 <BRO>
A;Cross-references: GB:L34075; NID:g508481; PIDN:AAA58486.1; PID:g508482
C;Genetics:
A;Gene: GDB:FRAP1; FRAP; RAFT1
A;Cross-references: GDB:597698; OMIM:601231
A;Map position: 1p36-1p36
C;Superfamily: yeast TOR2 protein

Query Match   100.0%; Score 9413; DB 2; Length 2549;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKDPDPDENPGVINNVLA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741 LEHSGIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILKDPDPDENPGVINNVLA 800

Qy     61 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQIVASTGYVVEPYRK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801 TIGELAQVSGLEMRKWVDELFIIIMDMLQDSSLLAKRQVALWTLGQIVASTGYVVEPYRK 860

Qy    121 YPTLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    861 YPTLEVLLNFLKTEQNGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES 920

Qy    181 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    921 KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK 980

Qy    241 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    981 SLGLKCVQFLPQVMPTFLNVIRVCDGAIREFLFQQLGMLVSFVKSHIRPYMDEIVTLMRE 1040

Qy    301 FWVMNTSIQSTIILLIEQIVVALGGEPFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 FWVMNTSIQSTIILLIEQIVVALGGEPFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI 1100

Qy    361 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV 1160
```

```
Qy  1501  PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVPEHAVNNTAGDDLA  1560
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2241  PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVPEHAVNNTAGDDLA  2300

Qy  1561  KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD  1620
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2301  KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGDRHPSNLMLDRLSGKILHIDFGD  2360

Qy  1621  CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  1680
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2361  CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  2420

Qy  1681  FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTVPESIH  1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2421  FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTVPESIH  2480

Qy  1741  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  1800
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2481  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  2540

Qy  1801  CYIGWCPFW  1809
               |||||||||
Db  2541  CYIGWCPFW  2549
```

RESULT 2
A54837
rapamycin/FKBP12 target 1 - rat
N;Alternate names: RAFT
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 24-Nov-1999
C;Accession: A54837
R;Sabatini, D.M.; Brdjument-Bromage, H.; Lui, M.; Tempst, P.; Snyder, S.H.
Cell 78, 35-43, 1994
A;Title: RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashio
A;Reference number: A54837; MUID:94306515; PMID:7518356
A;Accession: A54837
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2549 <SAB>
A;Cross-references: GB:U11681; NID:g511228; PIDN:AAA20091.1; PID:g511229
C;Superfamily: yeast TOR2 protein

```
Query Match            99.2%;   Score 9341;   DB 2;    Length 2549;
Best Local Similarity   99.1%;   Pred. No. 0;
Matches 1793;    Conservative  5;   Mismatches  11;   Indels  0;   Gaps  0;

Qy     1  LEHSSIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKLKDPPDPNPGVINNVLA  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741  LEHSSIGRIKEQSARMLGHLVSNAPRLIRPYMEPILKALILLKLKDPPDPNPGVINNVLA  800

Qy    61  TIGELAQVSGLEMRKKVDELFIIMDMLQDSSLLAKRQVALWTLGQLVASTGYYVEPYRK  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801  TIGELAQVSGLEMRKKVDELFVIMDMLQDSSLLAKRQVALWTLGQLVASTGYYVEPYRK  860

Qy   121  YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861  YPTLLEVLLNFLKTEQNQGTRREAIRVLGLLGALDPYKHKVNIGMIDQSRDASAVSLSES  920

Qy   181  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK  240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921  KSSQDSSDYSTSEMLVNMGNLPLDEFYPAVSMVALMRIFRDQSLSHHHTMVVQAITFIFK  980

Qy   241  SLGLKCVQFLPQVMPTFLNVIRVCDGAIREPLFQQLGMLVSFVKSHIRPYMDEIVTLMRE  300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981  SLGLKCVQFLPQVMPTFLNVIRVCDGAIREPLFQQLGMLVSFVKSHIRPYMDEIVTLMRE  1040

Qy   301  FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSPGRIVSIKLLAAI  360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1041  FWVMNTSIQSTIILLIEQIVVALGGEFKLYLPQLIPHMLRVFMHDNSQGRIVSIKLLAAI  1100

Qy   361  QLFGANLDDYLHLLLPPIVKLFDAPEAPLPSRKAALETVDRLTESLDFTDYASRIIHPIV  420
```

```
Qy   421  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVXGY  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1161  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVXGY  1220

Qy   481  TLADEEDPLIYQHRMLRSGCGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1221  TLADEEEDPLIYQHRMLRSGGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW  1280

Qy   541  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1281  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  1340

Qy   601  IELALTSQDIAEVTQTLLNLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1341  IEIALTSQDIAEVTQTLNLLAEFMEHSDKGPLPLRDDNGIVLLGERAAKCRAYAKALHYK  1400

Qy   661  ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1401  ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL  1460

Qy   721  VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1461  VAYDKKMDTNKDDPELMLGRMRCLEALGRWGQLHQQCCEKWTLVNDETQAKMARMAAAAA  1520

Qy   781  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLAHQDLFSLAQQCIDKARDLLDAELTAMAG  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1521  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAG  1580

Qy   841  ESYSRAYGAMVSCHMLSELEBVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR  900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1581  ESYSRAYGAMVSCHMLSELEBVIQYKLVPERREIIRQIWERLQGCQRIVEDWQKILMVR  1640

Qy   901  SLVVSPHEDMRTWLKYASLCGKSGRIALALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAY  960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1641  SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTIVLLLGVDPSRQLDHPLPTVHPQVTYAY  1700

Qy   961  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDDQQEKQELHKLMARCFLKLGEWQLN  1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1701  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDDQQEHKQELHKLMARCFLKLGEWQLN  1760

Qy  1021  LQGINESTIPKVLQYYSAATHDRSWYKAWHAWAVMNFEAVLIYKHQNQARDEKKKLRHA  1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1761  LQGINESTIPKVLQYYSAATHDRSWYKAWHAWAVNFEAVLIYKHQNQARDEKKKLRHA  1820

Qy  1081  SGANITNAITTAATATTAATATTTASTEGSNSESEABSTENSPTPSPLQKKVTEDLSKTLLMY  1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1821  SGANITNAITAATTAATATTTASTEGSNSESEABSS'ENSPTPSPLQKKVTEDLSKTLLMY  1880

Qy  1141  TVPAVQGFFRSISLSRGNNLQDTLRVLTIWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1881  TVPAVQGFFRSISLSRGNNLQDTLRVTIWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP  1940

Qy  1201  QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALLYPLTVASKSTTARHNAANKLIKNMCE  1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1941  QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALLYPLTVASKSTTARHNAANKLIKNMCE  2000

Qy  1261  HSNTLVQQAMMVSEELIRVALLWHEMWHEGLEBASRLYFGERNVKGMFEVLEPLHAMMER  1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2001  HSNTLVQQAMMVSELIRVALLWHEMWHEGLEBASRLYFGERNVKGMFEVLEPLHAMMER  2060

Qy  1321  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS  1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2061  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTS  2120

Qy  1381  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH  1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2121  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH  2180

Qy  1441  EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV  1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2181  EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV  2240
```

```
1501  Qy  PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDLA  1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2241  Db  PHCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTIMQKVEVFEHAVNNTAGDDLA  2300

1561  Qy  KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGD  1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2301  Db  KLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYIILGLGDRHPSNLMLDRLSGKILHIDFGD  2360

1621  Qy  CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2361  Db  CFEVAMTREKFPEKIPFRLTRMLTNAMEVTGLDGNYRITCHTVMEVLREHKDSVMAVLEA  2420

1681  Qy  FVYDPLLNWRLMDTNTKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  1740
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
2421  Db  FVYDPLLNWRLMDTNAKGNKRSRTRTDSYSAGQSVEILDGVELGEPAHKKTGTTVPESIH  2480

1741  Qy  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  1800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2481  Db  SFIGDGLVKPEALNKKAIQIINRVRDKLTGRDFSHDDTLDVPTQVELLIKQATSHENLCQ  2540

1801  Qy  CYIGWCPFW  1809
            |||||||||
2541  Db  CYIGWCPFW  2549
```

```
1101  Db  QLFGANLDDYLHLLLPPIVKLFDAPEVPLPSRKAALETVDRLTESLDFTYASRIIIHPIV  1160

 421  Qy  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1161  Db  RTLDQSPELRSTAMDTLSSLVFQLGKKYQIFIPMVNKVLVRHRINHQRYDVLICRIVKGY  1220

 481  Qy  TLADEEDPLIYQHRMLRSGQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW   540
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
1221  Db  TLADEEDPLIYQHRMLRSSQGDALASGPVETGPMKKLHVSTINLQKAWGAARRVSKDDW   1280

 541  Qy  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1281  Db  LEWLRRLSLELLKDSSSPSLRSCWALAQAYNPMARDLFNAAFVSCWSELNEDQQDELIRS  1340

 601  Qy  IELALTSQDIAEVTCQTLLNLAEFMEHSDKGPLPLRDNGIVLLGERAAKCRAYAKALHYK  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1341  Db  IELALTSQDIAEVTCQTLLNLAEFMEHSDKGPLPLRDNGIVLLGERAAKCRAYAKALHYK  1400

 661  Qy  ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMCHFGELEIQATWYEKLHEWEDAL  720
            ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
1401  Db  ELEFQKGPTPAILESLISINNKLQQPEAAAGVLEYAMKHFGELEIQATWYEKLHEWEDAL  1460

 721  Qy  VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKMARMAAAAA  780
            |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
1461  Db  VAYDKKMDTNKDDPELMLGRMRCLEALGEWGQLHQQCCEKWTLVNDETQAKVARMAAAAA  1520

 781  Qy  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVIALHQDLFSLAQQCIDKARDLLDAELTAMAG  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1521  Db  WGLGQWDSMEEYTCMIPRDTHDGAFYRAVIALHQDLFSLAQQCIDKARDLLDAELTAMAG  1580

 841  Qy  ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCCRIVEDWQKILMVR  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1581  Db  ESYSRAYGAMVSCHMLSELEEVIQYKLVPERREIIRQIWWERLQGCCRIVEDWQKILMVR  1640

 901  Qy  SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1641  Db  SLVVSPHEDMRTWLKYASLCGKSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAY  1700

 961  Qy  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1701  Db  MKNMWKSARKIDAFQHMQHFVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLN  1760

1021  Qy  LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYHQNQARDEKKKLRHA   1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1761  Db  LQGINESTIPKVLQYYSAATEHDRSWYKAWHAWAVMNFEAVLHYHQNQARDEKKKLRHA   1820

1081  Qy  SGANITNATTAATTA STEGSNSESEAASTENSPTPSPLQKKVTEDLSKTLLMY        1140
            ||||||||||  ||| :  ||||||||||||||:  |||||||||||||||||:
1821  Db  SGANITNATTTATTAASAAAATSTEGSNSESEAASNESSPTPSPLQKKVTEDLSKTLLLY  1880

1141  Qy  TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1881  Db  TVPAVQGFFRSISLSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIP  1940

1201  Qy  QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE   1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1941  Db  QLIARIDTPRPLVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTARHNAANKILKNMCE   2000

1261  Qy  HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER  1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2001  Db  HSNTLVQQAMMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER  2060

1321  Qy  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQWDLYYHVFRRISKQLPQLTS   1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2061  Db  GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQWDLYYHVFRRISKQLPQLTS   2120

1381  Qy  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH  1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2121  Db  LELQYVSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGH  2180

1441  Qy  EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV  1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2181  Db  EFVFLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWV  2240
```